Statistical Sample Size Determination Methods for Inspections of Engineering Systems

by

Eishiro Higo

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Terje Haukaas
Professor, Dept. of Civil Engineering,
University of British Colombia

Supervisor(s): Mahesh D. Pandey
Professor, Dept. of Civil and Environmental Engineering,
University of Waterloo

Internal Member: Giovanni Cascante
Professor, Dept. of Civil and Environmental Engineering,
University of Waterloo

Internal Member: Wei-Chau Xie
Professor, Dept. of Civil and Environmental Engineering,
University of Waterloo

Internal-External Member: Keith W. Hipel
University Professor, Dept. of Systems Design Engineering,
University of Waterloo

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A statistical sample size determination (SSD) method is designed for the maintenance of engineering components of similar structure within an overall system. The maintenance problem is defined as a sequential decision-making process, in which the optimal sample sizes are derived by an approach based on the value of information (VoI) concept.

Firstly, various sample size determination methods are summarized, and their advantages and disadvantages are discussed. This comparison highlights that, in many cases, the VoI-based approach is superior to traditionally used methods. Existing standards for engineering components are then categorized, based on the comparison, and the rationale behind each standard is described. Potential advantages of using a VoI-based approach are suggested and discussed.

Secondly, the theoretical superiority of VoI-based methods is demonstrated in the context of a diagnostic inspection problem, in which the traditional SSD method, the hypothesis-testing approach, can be defined. After the hypothesis-testing context is translated into a sequential decision-making problem, theoretical and numerical results are compared for the VoI-based and traditional methods.

Thirdly, the models for condition-based maintenance problems are defined with a time-dependent degradation process called the gamma process. The models mathematically describe how temporal and parameter uncertainties of the degradation process affect VoI-based analysis. Computational calculation techniques are introduced and compared with each other. Additionally, the model is generalized as a dynamic programming problem and formulated as a multiple-inspection problem.

Finally, the effectiveness of the SSD approach is demonstrated through application to an

actual degrading system. Based on data from nuclear power plants, numerical analyses are shown for both single and two inspection cases. The results provide operators with guidelines for maintenance and inspection policies that minimize the expected cost throughout the remaining lifetime of the system.

Acknowledgements

My greatest thanks go to my supervisor, Professor Mahesh D. Pandey, for his articulate suggestions, cheerful comments, generous patience, and financial support throughout my doctorate program. Without his help, the thesis could not have been completed. I am really honoured to have Professor Terje Haukaas from the University of British Colombia as my external examiner. My sincere gratitude is also due to Professor Keith W. Hipel, for his encouragement and helpful support for both my academic and personal lives in Canada. I would also like to thank my PhD internal committee members, Professor Giovanni Cascante and Professor Wei-Chau Xie.

I owe gratitude to Dr. Xianxun Yuan from Ryerson University, for the enlightening discussions and insightful comments that significantly enhanced the quality of my research. I would like to convey my appreciation to Dr. Norio Okada for giving me philosophical but, at the same time, practical messages when I stood at a crossroads in my life. Thanks are also due to Professor Liping Fang of Ryerson University for his special guidance on successful publishing, including techniques of professional writing and prompt responses to reviewers, and for holding me to a strict standard of work. Professor Kei Fukuyama from Tottori University in Japan has been a great mentor during my research life in Canada, and Dr. Mikko Jyrkama has shared his detailed knowledge about operating nuclear reactors. Sincere thanks to both. I would like to express my special thanks to Ms. Mary McPherson not only for her support for my writing but also for her cheerful comments on my study.

Moreover, I would like to acknowledge my dear friends, Dr. Sigong Zhang, Dr. Yingjun Deng, Dr. Noldanierick Manzana, Dr. Zhen Cai, Dr. Bin Liu, Mr. Sreehari Ramachandra, Mr. Indranil Hazra, Dr. Geourge Balomenos, and Mr. Conrad Hipel for their continuous support.

Finally, I thank my wife, Suhyun, my daughter, Naomi, and my parents, Shigeyoshi and Junko, for their love, support, and sacrifice throughout the long journey on this winding road.

Dedication

To Suhyun, Naomi, and my parents

Table of Contents

List of Tables

List of Figures

List of Algorithms

List of Abbreviations

ACC	Average coverage criterion
ALC	Average length criterion
AQL	Acceptable quality limit
DM	Decision maker
ENGS	Expected net gain of sampling
EVPI	Expected value of perfect information
EVPPI	Expected value of partial perfect information
EVSI	Expected value of sampling information
FAC	Flow-accelerated corrosion
HPD	Highest posterior density
HI	Histogram intersection
KL	Kullback-Leibler
LQ	Limiting quality
LQI	Life quality index
LTPD	Lot tolerable percent defective
MCMC	Markov chain Monte Carlo
MCS	Monte Carlo simulation
MDP	Markov decision process

MLE	Maximum likelihood estimation
OCC	Operating characteristic curve
POMDP	Partially observable Markov decision process
PWR	Pressurized water reactor
RQL	Rejectable quality level
SSD	Sample size determination
VoI	Value of information
WOC	Worst outcome criterion
WTP	Willingness to pay

List of Notation

a_i	ith option of replacement actions
a^o	Optimal action
C_F	Failure cost per component
C_I	Cost for inspection per component
C_P	Cost for preventive replacement per component
d	Width of the grey region
e_i	ith option of inspections
$\mathbb{E}_X\left[C(x)\right]$	Expectation of a function $C(x)$ with respect to a random variable X
$f_X(x)$	PDF of a random variable X
$F_X(x)$	CDF of a random variable X
$ga(x; a, b)$	PDF of a gamma random variable with shape parameter a and scale parameter b
H_0	Null hypothesis
H_1	Alternative hypothesis
$Iga(\mu; \alpha, \beta)$	PDF of a inverse-gamma ranadom variable with shape parameter α and scale parameter β
$L(x \mid Z = z)$	Likelihood function of x given $Z = z$
N	Population size

n	Sample size
$N(\mu,\ \sigma^2)$	Normal distribution with a mean μ and variance σ^2
$\mathbb{P}\left[X_2 > \rho_F \mid X_1 < \rho_F\right]$	Probability of observing $X_2 > \rho_F$ given $X_1 < \rho_F$
x_b	The break-even value of X at which two actions are equally preferred
x_{cr}	Critical value of X in hypothesis-testing approach
$\bar{x}$	Mean value of X
$\Gamma(z)$	Gamma function defined by $Gamma(z) = \int_0^\infty u^{z-1} e^u du$
ρ_F	Failure threshold

Chapter 1

Introduction

As our society matures, the cost of sustaining the built public infrastructure, such as roadways, bridges, and power plants, becomes enormous. Maintenance techniques for operating repairable engineering systems are now essential, and effective maintenance strategies are pursued for better safety and economic benefits. Although a large budget is allocated for infrastructure investment, still a significant funding gap exists between what is actually needed and the available funding. For example, according to the American Society of Civil Engineers (2016), federal, state, and local governments in the US will fund only 57% of the budget required for 2016 to 2025, which is an estimated 3.3 trillion U.S. dollars. In addition to promoting political efforts to increase the budget, authorities need to make the cost for infrastructure maintenance as low as possible, but also compatible with component safety.

In the maintenance of engineering components, which components and when to replace them are the main concerns and have been addressed by using many optimization approaches. Maintenance strategies can be roughly classified into two groups: time-based maintenance and condition-based maintenance. With time-based maintenance, which is sometimes called age-based replacement, components will be replaced at their scheduled times. This policy can be understood as an optimized strategy without inspection. Operators make plans based not on individual degradation processes, but on a general tendency. Thus, with a time-based maintenance policy, operators risk replacing components that still have a long remaining lifetime, which is wasteful. With condition-based maintenance, on the other hand, periodic inspection is scheduled, and decisions to replace a component or not are based on its inspected condition. Although time-based maintenance has been the majority choice for decades because of its easier implementation, condition-based maintenance has recently received attention as monitoring technologies develop. Pandey et al. (2009) reveal that condition-based maintenance is preferable when the uncertainty associated with degradation is relatively small, although in some cases time-based maintenance is better than condition-based maintenance. Most condition-based maintenance optimization studies propose determining the replacement criteria and inspection intervals needed for preventive maintenance (PM) based on the assumption that the conditions of all components are observed at each periodic inspection.

Nuclear power plants are one of the most critical infrastructures to be operated under a need to balance considerations of safety with those of generation efficiency. In order to remain operating, plants have to ensure authorities that they satisfy regulatory standards, and so must undergo inspections by an independent regulator. Sustained and efficient generation is desired by operating corporations, as it drives profit. To maintain safety and generation efficiency at high levels simultaneously, each component of a plant is regularly inspected. Usually, a nuclear power plant has a planned maintenance outage for two to eight weeks every two years (Garland, 2014). By inspecting components, operators can replace only those that are faulty or unlikely to satisfy required performance until the next inspection outage. Since a nuclear power plant consists of many sub systems, which comprise a number of components, the probability of failure for each component should be kept low. Otherwise, the whole system would need to be shut down often because even failure of one component may affect and stop the whole system. To avoid this frustrating situation, during a maintenance outage, components are inspected and those that have a higher probability of failure are replaced immediately. In terms of safety, the larger the sample size inspected, the less the uncertainty about current and future states of the components.

However, inspections of nuclear power plant components usually become expensive because of several difficulties, such as the high-radiation area and the large number of components. For example, a 600 MWe (electric) CANDU reactor core has 380 fuel channels, which are pressure tubes (see Figure 1.1 (Garland, 2014)). Once the wall thickness of a single pipe drops below a set threshold, heavy water leakage can occur. The actual strategy in a maintenance outage is to inspect only a part of all components and estimate the others' conditions from the newest inspection outcomes and previous data. Thus, sample size determination becomes an important problem in balancing the safety or generating efficiency requirement with management cost.

1.1 Research Motivation

Standards and guidelines for sample size determination (SSD) have been published and used at actual sites, but they are missing theoretical rationale or rely on methods that ignore inspection cost. For example, the minimum sample sizes for components of CANDU reactors are summarized as guidelines (National Standards of Canada, 2014). Although the guidelines have worked well at actual operating sites, no theoretical rationale exists for the sample sizes and requirements for sample selection. Several other standards such as National Standards of Canada (2014) rely on traditional SSD methods, but these methods focus only on safety and cannot include cost-effectiveness in the inspection policy.

Except for the traditional SSD techniques, few SSD methods have been developed for condition-based maintenance. Since condition-based maintenance is based on data obtained from

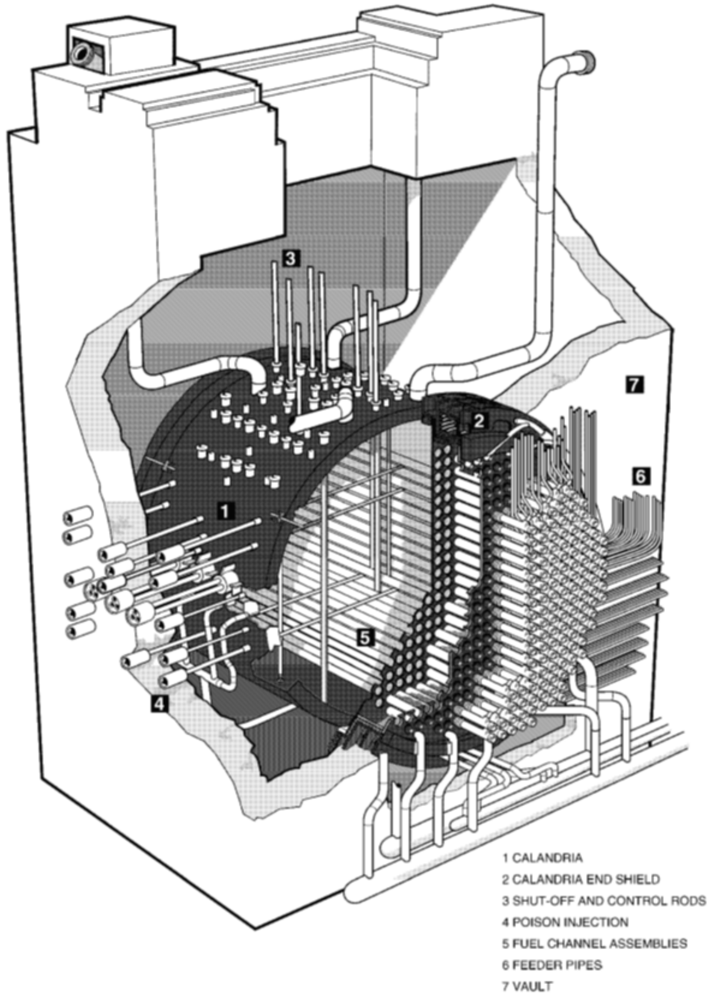

Figure 1.1: Reactor core of a CANDU 600 (Garland, 2014)

periodic inspection, SSD is ignored in most studies. The assumption of full-inspection needs to be revisited to include the inspection problem. As the example points out, there are cases like nuclear power plants where the sample size of each inspection needs to be considered in the condition-based maintenance strategy.

An SSD approach, proposed by Raiffa and Schlaifer (1961), called the value of information (VoI) concept, is gaining attention in structural health monitoring (Faber and Sorensen, 2002; Straub and Faber, 2004a,b; Bensi, 2010; Pozzi and Der Kiureghian, 2011; Straub, 2014; Memarzadeh and Pozzi, 2016; Konakli et al., 2016). It is used for calculating the benefit of obtaining information, and with it, operators can find a reasonable balance between the cost of inspection and the pay-off of the results. Despite noteworthy contributions of these studies, several limitations remain. First, these studies still have difficulty in dealing with system-level problems. Thus, SSD cannot be carried out with the existing VoI-based analyses. Second, these studies are not focused on applying the method to deterioration models and so have been applied to realistic degradation process models and do not explain how to extend the method to those degradation models. Third, they fail to show how much the VoI-based method can reduce the expected cost compared with traditional SSD methods. These studies, relying on Bayesian statistics, use different terminologies from frequentist statistics, making comparisons difficult.

As another approach for an inspection decision problem, the partially observable Markov decision process (POMDP) has been developed for condition-based maintenance in recent years (Papakonstantinou and Shinozuka, 2014a,b,c; Memarzadeh and Pozzi, 2016). The approach focuses on measurement errors and optimizes not only maintenance actions but also inspection policy. It successfully generalizes multiple-inspection problems by simplifying the states of a system, observation outcomes, and maintenance/inspection actions as discrete values. Schöbi and Chatzi (2016) extend the POMDP for continuous-state problems. However, it still cannot deal with the SSD problem in the analysis scheme because it has focused only on a single-component system. Moreover, it has not considered other types of uncertainties, such as parameter and temporal uncertainties.

1.2 Objectives

The overall goal of this thesis is to develop widely applicable SSD methods for investigating system-level condition-based maintenance problems, and thus enhance decision-making for engineering-component maintenance. In particular, the thesis covers the following questions and approaches to solving them:

- What are the basic characteristics of SSD methods based on VoI concept within the context of engineering-component maintenance problems? The author develops a simple model that represents the maintenance problem and shows how the VoI-based SSD methods work;
- How do the proposed SSD methods differ from frequentist techniques? The author compares these two approaches and shows the strengths and weaknesses of each;
- How can a system-level maintenance problem be defined for a time-dependent degradation process? The author applies a stochastic degradation process modelled as a gamma process, with which one can include temporal and parameter uncertainties. Because of the nature of the process, one can describe all possible system conditions in a simple manner. Note that the system is defined as a group of homogeneous components whose degradation levels can be treated as independent and identically distributed random variables, although they become dependent once their distribution parameters share a common parameter. The system is repairable, and each component can be replaced;
- How can an SSD method be applied to multiple-component multiple-inspection problems?

The author explains how to combine the SSD method with dynamic programming. As an example, the VoI-based method is demonstrated with data from a real operating system.

1.3 Thesis Overview

The thesis organization is illustrated in Figure 1.2. The thesis is organized in the following manner.

- Chapter 2 reviews the relevant literature on SSD methods, discusses the advantages and disadvantages of each technique, and confirms the superiority of the VoI-based SSD method.
- Chapter 3 provides an SSD analysis for diagnostic inspection of a component population. An inspection and replacement problem is defined and is combined with a binomial states model. This chapter provides hypothesis-testing-based and VoI-based methods for finite population cases, compares these two methods, and discusses their differences.
- In Chapter 4, the VoI-based SSD method is applied to a linear degradation process model, a random rate model. First, a two-stage decision-making problem is defined, and the VoI-based SSD method is formulated for both single-component and multiple-component system cases. A numerical example is demonstrated, and the characteristics of the method are analysed.
- Chapter 5 deals with a gamma process model, which is a time-dependent stochastic degradation, in maintenance problems with temporal uncertainties. It starts with a single-component case and develops it into a multiple-component system case. The author demonstrates the given approach on a realistic numerical example of maintenance of nuclear power plant components.
- Chapter 6 extends the model developed in Chapter 5 to a case with parameter uncertainty. A single-component case is introduced and used to define a multiple-component system case. The author also discusses the impacts of reducing different types of uncertainties: temporal and parameter uncertainties.

- Chapter 7 extends the proposed model to the two-inspection problem by formulating it as a dynamic programming situation. It derives a general model for multi-inspection problems and demonstrates a two-inspection problem with data from a real operating system.
- Chapter 8 summarizes the contributions of the research and points out avenues to follow in future work.

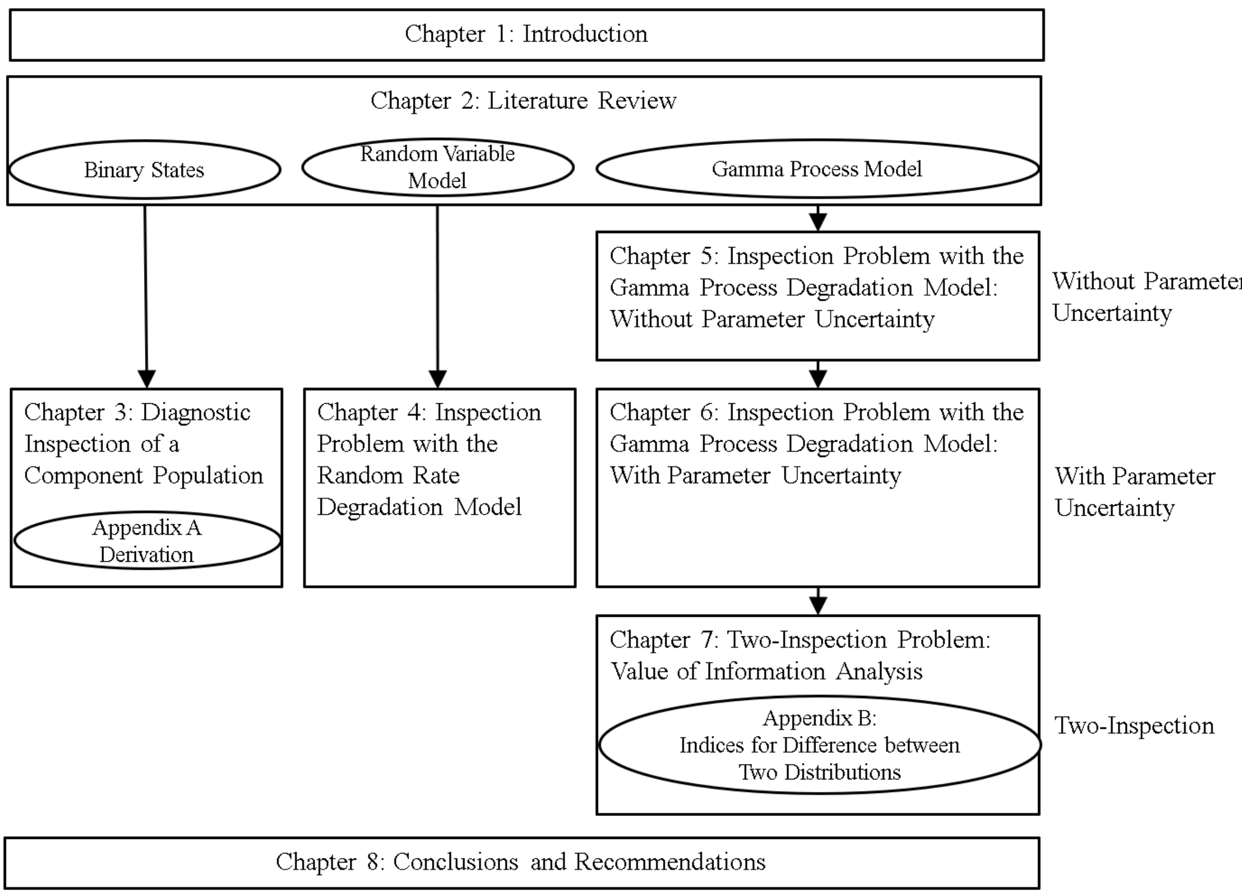

Figure 1.2: Graphical overview of the thesis

Chapter 2

Literature Review

This chapter reviews the relevant literature on sample size determination (SSD) methods. The SSD methods were originally developed as a part of statistical experimental design in the early twentieth century. Since sample size is a key factor of frequentist statistic analysis such as hypothesis-testing and parameter estimation, as a part of designing these analyses, a variety of SSD methods were derived. These methods have been widely used for a long time and are still in the major approach of SSD. Although it is still in the minority, another SSD method, VoI-based method, has been getting attention over the last decade. These methods have developed independently and have never been compared with one another. This chapter summarises these methods and builds a basis for comparison.

2.1 Uncertainties and Random Variables in Degradation Processes

Inspection data contribute to reducing uncertainties about component deterioration. The uncertainties can be categorized into three types: measurement error, temporal uncertainty, and random effects (Verbeke and Molenberghs, 2009; Yuan, 2007). In addition to these three, in the context of Bayesian statistics, we need to include parameter uncertainty, which is an uncertainty in a underlying mathematical model. First, measurement error is a gap between observed data and the true state of the same component at the same time. This uncertainty is dominant when we estimate the current states of inspected pipes. Second, temporal uncertainty means uncertainty for a future state. Even though we know the true current state, future states cannot be predicted exactly since the degradation process is stochastic. This temporal uncertainty needs to be considered when estimating the future states of inspected pipes. Third, random effects represent the heterogeneity of a pipe compared with other pipes that should have the same characteristics theoretically. Random effects appear in problems estimating other pipes based on the data of already-inspected pipes. Parameter uncertainty represents imperfect information about population parameters, which are treated as random variables in the context.

Uncertainties can be classified into two groups: aleatory and epistemic. Aleatory uncertainties arise from natural or unpredictable variation and are in general not reducible, whereas epistemic uncertainties are from lack of knowledge about the focusing random variables and can be reduced by increasing inspection accuracy and/or the size of sampling inspection. Measurement error and parameter uncertainty are classified as epistemic uncertainties, and temporal uncertainty and random effect are aleatory uncertainties.

Under the context of inspection planning, operators can "reduce" temporal uncertainties by planning a new inspection at a future time. Adding another decision-making time and inspecting components, the operators can reduce the time interval in which they need to forecast the degradation process of the components. In the inspection planning, both the aleatory and epistemic uncertainties need to be considered.

2.2 Degradation Process

The object of inspection is to observe the condition of a component that is degrading during system operation. Obtained outcomes update information about the degradation process and consequently contribute to better prediction of future conditions. For the prediction, operators select and use a mathematical degradation model, which is essential for maintenance planning.

The overall classification of degradation models is illustrated in Figure 2.1. Probabilistic degradation models can be generally classified into two groups: random variable and stochastic process models. The random variable models assume that the randomness exists among components, but the path of degradation process is deterministic. The randomness is represented by vector of random variables. Thus, they do not include aleatory uncertainties so that the future condition can be precisely predicted if no measurement errors exist in the observation of current contritions. The stochastic process models include temporal uncertainties, which are aleatory uncertainties, and are associated with progression of degradation over time. The process itself includes uncertainties so that a future is still uncertain even if the current state is observed without measurement errors. Note that the models satisfy the Markov property as long as they assume independent increments, which is more restrictive than the Markov property (van Noortwijk, 2009).

The stochastic process models can be split into two subgroups based on whether they assume discrete or continuous states. Discrete-state models are classified as discrete-time Markov processes, which discretize time and consequently have discrete-states. The deterioration progression is modelled as transitions between states that are defined by probability matrix. The models are usually used in the Markov decision process (MDP) or its derivation, the partially observable Markov decision process (POMDP). The continuous-state models, which are identical to continuous-time Markov processes, are vary, such as the gamma process, the Wiener process

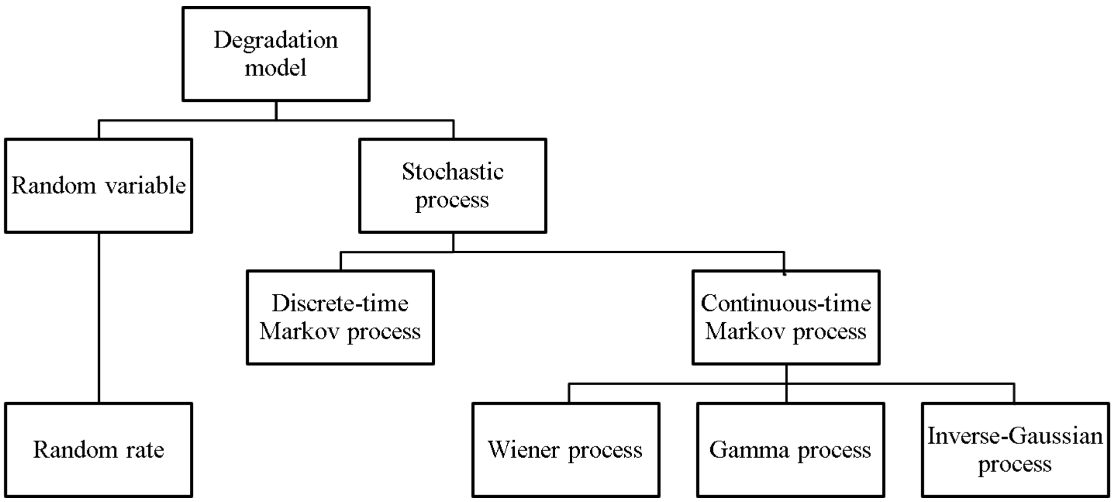

Figure 2.1: Classification of degradation models

(also called the Gaussian process or the Brownian motion with drift,) and the inverse-Gaussian process (Ye et al., 2015).

A gamma process is a continuous-time Markov process with stationary, independent, and gamma distributed increments (Abdel-Hameed, 1975). With a gamma process, the distribution of a future state follows a gamma distribution wherein one of the parameters is proportional to a time interval between the two inspection timings. Because of their simple mathematical form and memoryless property, gamma processes have been used in modelling a variety of degradation phenomena (Yuan, 2007). Large applicability of gamma processes has been shown, including for diverse materials and failure modes, such as sand nourishment erosion (van Noortwijk and Peerbolte, 2000), rock rubble displacement in sea bed protection (van Noortwijk et al., 1995), concrete creep (Cinlar et al., 1977), scour-hole development on concrete surface (van Noortwijk and Klatter, 1999), corrosion of carbon steel pressure vessels (Kallen and van Noortwijk, 2005), fatigue crack growth (Lawless and Crowder, 2004), feeder wall thinning corrosion (Yuan et al., 2008), and diameter expansion of fuel channels (Yuan et al., 2006).

When the increments follow a normal distribution, the continuous-time Markov process is the Wiener process (van Noortwijk, 2009). Although the Brownian motion, with which the increments can be negative values, is inadequate for modelling the deterioration process, because of mathematical advantages, the Wiener process has been widely used in a wide range of applications such as bridge beam degradation (Wang, 2010) and magnetic head wearing (Ye et al., 2013).

Similar to the gamma process, the inverse-Gaussian process is a monotonically increasing degradation process with independent and inverse-Gaussian distributed increments. The process is flexible for modelling heterogeneous degradation of systems because of its easiness to include random effects. The inverse-Gaussian process was first introduced by Wang and Xu (2010) and has gained attention recently. The usefulness of the process for condition-based maintenance, especially for heterogeneous system, has been investigated by Qin et al. (2013), Ye and Chen (2014), Chen et al. (2015), and Peng et al. (2017).

2.3 Value of Information Analysis

The value of information concept is defined in Bayesian decision analysis. Bayesian decision analysis can be understood as a branch of statistical decision theory. Building on game theory, which was originally proposed by von Neumann and Morgenstern (1947), Wald (1950) initiated and developed the statistical decision theory. In the theory, a decision maker plays a game against an opponent, "nature." Nature controls randomness in the decision problem, and the decision maker takes an action through calculating possible consequences and their probability of occurrence. Since then, statistical decision theory has been introduced, and extended by many researchers, such as Blackwell and Girshick (1954), Chernoff and Moses (1959), Ferguson (1967), Hadley (1967), Weiss (1961), and DeGroot (1970). In certain of these studies, Bayesian decision analysis, including the VoI concept, was proposed, by Pratt et al. (1995), Raiffa and Schlaifer (1961), and Schlaifer (1959). Since then Bayesian decision analysis has been applied to many fields, including civil engineering problems; one initial study of civil engineering applications was produced by Benjamin and Cornell (1970). The use of VoI and its derivations has become frequent over this last decade, especialy in healthcare science (Steuten et al., 2013).

VoI is a value for a given observation result, and its expected value with respect to as-yet-unknown obtaining inspection outcomes is called the expected value of sample information (EVSI). The EVSI is a function of sample size and represents by how much the decision maker benefits from the observed data because that data reduces uncertainty about a component's true state. The $EVSI$ has the following attributes:

- is an expected value of the VoI;
- treats the observation outcomes as probabilistic variables;
- is a function of one or more parameters of inspection, such as sample size;
- identifies an operator's expected benefit from observed data.

As a utility function against sample size, the expected net gain of sampling (ENGS) is defined as EVSI minus observation cost. Maximizing ENGS, we can determine the optimal sample size.

2.3.1 Random Variables and Sets of Options

Typical Bayesian pre-posterior decision analysis is shown in Figure 2.2, which pictures the whole decision process, relations among the four key variables, and consequences for each possible set of four values. The terminal cost $C(e, z, a, x)$ is determined by four values (e, z, a, x) from four data sets: the set of possible experiments (E), the set of potential outcomes of all experiments in the set (Z), the set of possible terminal acts (A), and the set of possible "states of the world"

(X). Note that E and A are the set of possible options for a decision maker, and Z and X are treated as random variables in this study. Using z (results of e), the decision maker updates the estimate of x and chooses an action, a, because it maximizes the expected terminal cost. Even though the true state is pre-determined by chance, the node for x comes last because the decision maker is only able to know the true value through a consequence of his/her decision.

In Bayesian pre-posterior analysis, the terms "prior" and "posterior" mean before and after experiments, respectively, and "pre-posterior" is used when we are considering a posterior situation but are actually still in a prior situation. Bayesian pre-posterior decision analysis is an optimization of the whole decision problem from the perspective of a decision maker before observations. In detail, this analysis is done to find the options of e and a that minimize the expected cost related to unknown values z and x.

2.3.2 Cost Function

The cost function can be separated into two parts as follows:

$$C(e, z, a, x) = C(a, x) + Cost(e, z), \tag{2.1}$$

where $C(a, x)$ is the re-defined "cost function," which is used in most VoI analysis and also used in Chapter 3; and $Cost(e, z)$ is the total inspection cost; for instance, a simple model is a linear function of a sample size, n, as $Cost(e, z) = Cost(n) = n \cdot C_I$, where C_I is the cost for an observation. For later analysis, we also define "prior cost" as a cost without experiment and "posterior cost" as a cost in a case with a certain e and z, excluding the cost for the experiment; respectively, the two are represented as $C(a, x)$ and $C(a, x \mid e, z)$ in Figure 2.2. Within this Bayesian pre-posterior analysis, EVSI can be calculated with two different expected costs.

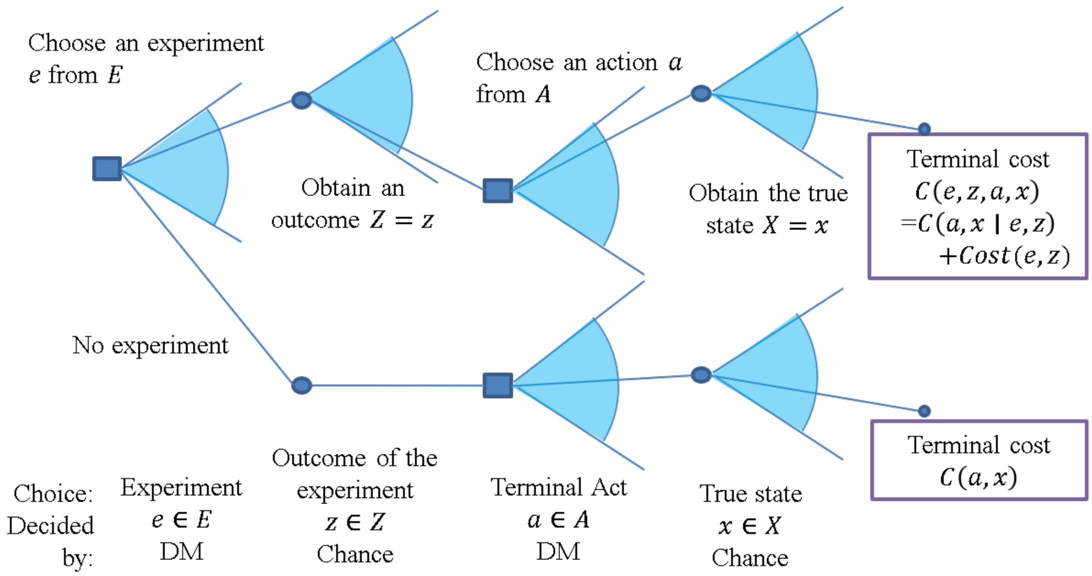

Figure 2.2: Extended decision tree for a Bayesian pre-posterior analysis

2.3.3 Prior Analysis

A prior analysis is the optimisation of the expected cost with respect to actions, A, without any new information. Decision makers evaluate the expected cost of each action based only on prior information, $\mathbb{E}_X\left[C(a, X)\right]$. The prior information is summarized as a probability density (or mass) function; this is called a prior distribution. There are no updates about unknown values, X. Through the analysis, decision makers obtain the best action and its expected consequence, $\min_a\{\mathbb{E}_X\left[C(a, X)\right]\}$, without sampling inspection.

2.3.4 Posterior Analysis

A posterior analysis is the optimization of expected cost with respect to A when decision makers have specific outcomes, such as $Z = z$, from a sampling inspection, $E = e$. Decision makers use z to update information about the unknown values, X. Combined the obtained outcomes with a prior distribution of X, a posterior distribution is derived. This updating can be calculated with the classical Bayes rule:

$$\mathbb{P}\left[X = x \mid Z = z\right] = \frac{\mathbb{P}\left[Z = z \mid X = x\right]\mathbb{P}\left[X = x\right]}{\mathbb{P}\left[Z = z\right]}, \tag{2.2}$$

where $\mathbb{P}\left[X = x \mid Z = z\right]$ is a posterior probability that a random variable X becomes a value x, which means that $X = x$; $\mathbb{P}\left[X = x\right]$ is a prior probability of $X = x$; $\mathbb{P}\left[Z = z\right]$ is a probability of the occurrence of $Z = z$; and $\mathbb{P}\left[Z = z \mid X = x\right]$ is the probability of the occurrence of $Z = z$ conditional on $X = x$, which is called a likelihood. If we treat only the observation $Z = z$ as fixed, $\mathbb{P}\left[Z = z\right]$ is also a constant value; we can then re-write Equation (2.2) as a function of x as follows:

$$f_{X|Z}(x \mid z) = L(x \mid Z = z) \cdot f_X(x) \cdot const, \tag{2.3}$$

where $f_{X|Z}(x \mid z)$ is a posterior distribution; $f_X(x)$ is a prior distribution; and $L(x \mid Z = z)$ is a likelihood function, equalling $\mathbb{P}\left[Z = z \mid x\right]$. Note that the likelihood function is $L(x \mid Z = z)$, not $l(Z = z \mid x)$, since it is a function of x, not $Z = z$, although the form is derived from $\mathbb{P}\left[Z = z \mid x\right]$. Decision makers can thus derive the best action and its expected consequence, $\min_a \mathbb{E}_{X|z}\left[C(a, X \mid e, z)\right]$.

2.3.5 Pre-posterior Analysis

Pre-posterior analysis is an optimization of the whole decision problem at the time of inspection planning. Decision makers can observe additional samples, thereby reducing his/her uncertainty of the true state, X, and want to minimize the expected cost with respect to inspection options, E. As the sample size increases, the expected cost decreases. By incorporating the as-yet-unknown sampling inspection outcome, we obtain the expected cost: $\mathbb{E}_Z \left[\min_a \mathbb{E}_{X|z} \left[C(a, X \mid e, z)\right]\right]$. This analysis determines the options of e and a that minimize expected costs on unknown values Z and X.

Comparison between the expected cost with and without the additional observation is called the Expected Value of Sample Information (*EVSI*); this comparison represents the additional expected benefit from the observations. *EVSI* is described as a function of a sample size, n, as follows:

$$EVSI(n) = \min_a \{\mathbb{E}_X \left[C(a, X)\right]\} - \mathbb{E}_Z \left[\min_a \mathbb{E}_{X|z} \left[C(a, X \mid e, z)\right]\right]. \tag{2.4}$$

Since *EVSI* is an expected value prior to additional observations, the second term is an expectation for both an uncertain event and the as-yet-unknown observations.

2.3.6 Expected Net Gain of Sampling (*ENGS*)

While a larger sample size helps to estimate the true state of the world, the observations involve more costs as the sample size increases. Extracting the observation cost from *EVSI*, we define another net benefit as the Expected Net Gain of Sampling (ENGS):

$$ENGS(n) = EVSI(n) - nC_I \tag{2.5}$$

where C_I is the cost for each sample observation. This *ENGS* is a function of sample size n, and we estimate the optimal sample size that maximizes the *ENGS*.

A more-detailed definition is as follows:

$$\begin{aligned} ENGS(n) &= \min_a\{\mathbb{E}_X\left[C(a,X)\right]\} - \mathbb{E}_Z\left[\min_a \mathbb{E}_{X|z}\left[C(a,X \mid e,z)\right]\right] \\ &= \min_a \int_{\mu\in M} C(a,x) f_X(x)dx \\ &\quad - \int_{z\in Z}\left[\int_{x\in X} \min_a\{C(a,x \mid n,s_n) f_{X|z}(x \mid e,z)dx\}\right] f_Z(z \mid z)dz, \end{aligned} \tag{2.6}$$

where $\mathbb{E}_{X|z}\left[g(X)\right]$ is an expectation of $g(X)$ on X, conditional on a given condition of $Z = z$; and $f_{X|z}(x \mid e,z)$ is a posterior probability density or mass function of a random variable, X, conditional on e and z. Since $ENGS$ is an expected value prior to new observations, the second term is an expectation for both an uncertain event and the as-yet-unknown observations. This $ENGS$ is a function of sampling inspection options, that is, n in the N component problem, and we estimate the optimal sample size that maximizes the $ENGS$.

2.4 Sample Size Determination Methodologies

SSD methods that are used for decision-making problems are classified here based on the ideas behind the methods. A typical separation criterion is whether the method uses a Bayesian approach or not. In statistics, researchers have argued for more than a hundred years, divided by their basic stances: frequentist or Bayesian (McGrayne, 2011). A frequentist treats probability as frequency after infinite trials, or a large sample size, and considers that unknown parameters are fixed values, and data is just an appearance of the value with randomness. On the other hand, a Bayesian assumes that given data is a set of definite values, and unknown parameters can only be estimated based on the data. A Bayesian treats probability as a subjective expression for an unknown value based on known information. In that sense, Bayesian statistics has an affinity for decision-making problems.

Based on the main idea behind them, SSD methods are classified into three groups, according to their principles: optimal sample size maximizing expected utility, minimum sample size satisfying given requirements (frequentist), and minimum sample size satisfying given requirements (Bayesian). The first group takes a Bayesian approach; the methods in this group use prior knowledge and new observations. The second group uses a frequentist approach; the methods do not consider prior information but use given safety criteria. The third group combines prior information with some safety criteria. Comparing these three groups, this chapter provides an overview of this research.

2.4.1 Value of Information Approach (Bayesian)

As a straightforward requirement for SSD, considering expectations of random variables, the utility maximization approach finds the optimal sample size that minimizes a total cost or maximizes a total value. This is a fully Bayesian approach, which is needed to determine a utility function (Adcock, 1997; Lindley, 1997). With a fully Bayesian method, sample size is determined by maximizing the expected utility; for example, Lindley (1997) demonstrates a method for maximizing a logarithmic utility function. The SSD method with VoI concept can be classified as a utility maximization approach.

Several studies have applied the VoI concept in structural health monitoring (Faber and Sorensen, 2002; Straub and Faber, 2004a,b; Bensi, 2010; Pozzi and Der Kiureghian, 2011; Straub, 2014; Memarzadeh and Pozzi, 2016; Konakli et al., 2016). For example, inspection optimization methods can combine a fatigue-crack-growth model with the VoI concept (Madsen, 1997; Straub, 2014). Straub and Faber (2004a) propose an approach to determine what percentage of the inspection should be performed with $EVSI$ analysis.

Despite the large potential of contributing to component-maintenance problems, the method is still not generalized for condition-based maintenance since its calculation is computationally expensive. Thus, a variety of simplified models have been proposed, whose major characteristics are summarized in Table 2.1. Most avoid complexities by simplifying the system to a single-component (Pozzi and Der Kiureghian, 2011; Straub, 2014; Konakli et al., 2016) and/or by simplifying degradation processes as descretized conditions with transition among them (Faber and Sorensen, 2002; Straub and Faber, 2004a,b; Memarzadeh and Pozzi, 2016).

The studies for single-component systems focus on optimizing the timing and threshold of maintenance actions by using inspection data. Konakli et al. (2016) have challenged the full-inspection assumption in condition-based maintenance by using a partially observable Markov decision process (POMDP) recently developed for condition-based maintenance (Papakonstantinou and Shinozuka, 2014a,b,c; Memarzadeh and Pozzi, 2016). The approach focuses on measurement errors and optimizes not only maintenance actions but also inspection policy. It successfully generalizes multiple-inspection problems by discretizing the values of system states, observation outcomes, and maintenance/inspection actions. However, the study still cannot deal with SSD problems in the analysis scheme because it focuses only on a single-component system. Moreover, it ignores other types of uncertainties, such as parameter and temporal ones.

Studies extending the VoI concept to multiple-component problems have struggled to overcome the computational cost of integrating all possible inspection outcomes, since the space of the outcome becomes n dimensional if we identify each component's outcome separately. To reduce the number of possible outcomes, these studies use discretized inspection outcomes. Several studies use binary inspection outcomes: failure detected or not (Faber and Sorensen, 2002; Straub and Faber, 2004a,b). Memarzadeh and Pozzi (2016) apply POMDP to a five-component inspection problem; however, they have not discussed how to find the optimal sample size.

Thus, few related studies focus on SSD. Straub and Faber (2005) proposed a method to reflect the dependency of inspection costs on the number of inspected hot spots, using *EVSI* as an evaluation criterion. The authors discuss how the number of inspected hot spots affects the *EVSI* but do not analyse their results in terms of SSD. Several other studies in inspection planning have calculated and analysed the *EVSI* or VoI in their decision-making problems, but they do not focus on finding a best sample size. For example, Madsen et al. (1986) combines a fatigue-crack-growth model with a failure-probability updating feature to produce an inspection optimization method. Straub (2014) also provides how to derive the *EVSI* with a fatigue-crack-growth model. The author compared the cases of one and two measurement(s) on a component and showed how the measurement error influences the results. In contrast to the physical models, Pozzi and Der Kiureghian (2011, 2012) apply the VoI concept to a linear degradation model, and the VoI for each possible posterior regression covariance matrix is calculated as a demonstration. Several studies apply VoI-based analysis for observation location planning (Krause, 2008; Yoshida, 2015). Yoshida (2015) proposes an optimizing method for determining inspection locations and a sample size, by means of a Gaussian random field. These location planning studies have potential uses in considering the correlation among different components, although this approach has not yet been discussed in the literature.

Despite implementation of VoI concept in various applicable fields, two limitations remain. First, they focus on finding the best investigation interval, not the best sample size. Second, they are not aimed at applying the method to a variety of deterioration models; how to apply the method to other cases is not clearly explained. Thus, SSD with VoI concept needs to be modelled and built in a simple form with degradation process models, within the engineering context.

Moreover, all the previous studies have focused on parameter uncertainties and/or

Table 2.1: Major characteristics of studies

Study	Number of components	Degradation model	Observation result	Decision points in time
Faber and Sorensen (2002)	Single	No growth (binomial states)	Discrete	Single
Straub and Faber (2004a)	Multiple	Random variable (lognormal)	Discrete	Multiple (two)
Straub and Faber (2004b)	Multiple	Random variable (crack growth)	Discrete	Multiple (two)
Pozzi and Der Kiureghian (2011)	Single	No growth (multinomial states)	Continuous	Single
Straub (2014)	Single	Random variable (crack growth)	Continuous	Multiple (two)
Mamarzadeh and Pozzi (2016)	Multiple	Markov process	Discrete	Multiple
Konakli *et al.* (2016)	Single	Random variable (linear)	Continuous	Single

measurement errors and not on temporal uncertainties in their models. As described in Section 2.1, under the context of inspection planning, even temporal uncertainties can be reduced by setting a new inspection time. The POMDP approach includes temporal uncertainties as the transition probabilities between each time step; however, no studies evaluates the value of reducing the temporal uncertainties.

Procedure for Sample Size Determination with Value of Information Concept

A simple maintenance problem within engineering context is illustrated in Figure 2.3. The procedure for SSD with VoI concept is as follows:

1. Prior analysis

   a Calculate the expected cost for each action, a.

   b Choose the best action, a^o.

2. Posterior analysis

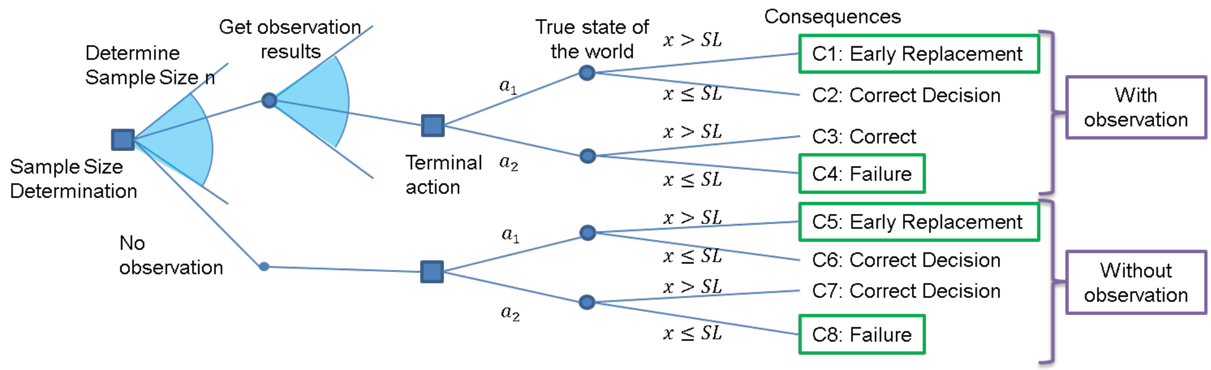

Figure 2.3: Extended decision tree for a maintenance problem as a Bayesian pre-posterior analysis

c Suppose an observation result, $Z = z$, is given, then calculate the expected cost for each action, a_z.

d Choose the best action, a_z^o, for each possible observation outcome, z.

3. Pre-posterior analysis

e Consider all possible observation outcomes $z \in Z$ and take the expected value of the cost with a_z^o, which is the optimal action conditional on z.

f Calculate $EVSI(n)$.

g Calculate $ENGS(n)$.

h Find the sample size n that maximizes the $ENGS(n)$.

2.4.2 Statistical Hypothesis-Testing Approach (Frequentist)

The main idea of the widely used hypothesis-testing approach is to derive a minimum sample size that satisfies required statistical error limits. For example, for given Type I and type II errors, and a gray region (where the two types of errors are not satisfied), a minimum required sample size has been derived by the U.S. Environmental Protection Agency (2006). Note that the type

I and II errors mean false positive and false negative errors. Details of this approach are further explained in Section 3.4. Comprehensive details of this approach are described by Cochran G. (1977), Kraemer and Thiemann (1988), and Desu and Raghavarao (1990). In this section, we introduce the hypothesis-testing approach with these criteria.

Note that we generalize the frequentist SSD approaches to "hypothesis-testing approach" although frequentist approaches can be used in two different contexts: hypothesis-testing and confidence interval estimation. The SSD methods in confidence interval estimation can be treated as an SSD in hypothesis testing without any restriction on type II errors, in terms of mathematical calculation. Because operators are concerned only with whether the condition satisfies required safety criteria in maintenance problems, we introduce only one-sided hypothesis testing cases.

Although these studies have been applied to several practical problems, there are two critical limitations. First, they cannot consider consequences resulting from sampling. inspection and maintenance costs do not affect the sample size. Second, this approach cannot combine prior information in a systematic manner.

With Required Threshold for Type I and II Errors

Suppose we need to estimate an unknown true value under two safety requirements for type I and II errors. The larger the sample size we observe, the more certainty we have for a decision based on the observed data. That is, the sample size increases, we have a smaller region in which the two safety requirements are not satisfied, which is called a "critical region" or "gray region." Once the maximum acceptable width of the critical region is given as the third criterion, we can determine the minimum sample size that satisfies the three restrictions (U.S. Environmental Protection Agency, 2006). Let x be the unknown true value. At this point, we do not have to consider how the random variable is distributed, but will return to this issue in a later chapter.

Let Y be a random variable that represents the observed sample value; as a typical example, we use sample mean. This random variable contains measurement errors, and is written as

$$Y = x + E, \tag{2.7}$$

where E follows a normal distribution, $N(\epsilon;\ 0,\ \sigma_\epsilon^2)$. We consider that Y follows normal distribution with the mean at x and the variance at σ_ϵ. The variance can be known or unknown; each case has a different formula for SSD. Respectively, the null hypothesis and alternative hypothesis are

$$\begin{aligned} H_0 :&\quad x \leq \rho_F \\ H_1 :&\quad x > \rho_F. \end{aligned} \tag{2.8}$$

We can set a safety criteria, α, for the maximum acceptable probability of rejecting H_0 in error when H_0 is true, which is called "type I error." Similarly, we can set β as the maximum acceptable probability, erroneously not rejecting H_0 when $x = x_b - d$, which means H_1 is true (a type II error). Let U denote the sample mean of the random variable Y. The sample mean also follows a normal distribution, with mean x and variance σ_ϵ^2/n.

The following subsections introduce sample size determination methods for known and unknown cases. We start from the known variance case as it is simpler, and increase the complexity by changing the assumption of known variance.

With known variance We assume that the variance of measurement error is known. We want to ensure that the estimation of the true value, x, is reasonably larger than a failure threshold,

ρ_F. We test $H_0 : x \leq \rho_F$ against $H_1 : x > \rho_F$ using the statistic

$$
\begin{aligned}
Z_0 &= \frac{U - \rho_F}{\frac{\sigma_\epsilon}{\sqrt{n}}} = \frac{U - x}{\frac{\sigma_\epsilon}{\sqrt{n}}} + \frac{x - \rho_F}{\frac{\sigma_\epsilon}{\sqrt{n}}} \\
&= Z + e(x), \qquad (2.9)
\end{aligned}
$$

where Z follows standardized normal distribution; and $e(x)$ is a function of x, which is defined as $e(x) \equiv \frac{x-\rho_F}{\sigma_\epsilon/\sqrt{n}}$. With the required probabilities of false positive, α, and false negative, β, the required minimum sample size is estimated as in the following procedure. We may want to find the minimum sample size such that

- The test hypothesis is falsely rejected with a probability of no greater than α, and
- Failure to reject the null hypothesis happens with a probability of no greater than β when the difference between the hypothetical population mean and the true population mean is as large as d.

The two requirements are summarized as an equation:

$$
\begin{aligned}
Pr\left[\text{Do not reject } H_0 \mid H_1 \text{ is true}\right] = \beta \quad &\Leftrightarrow \quad Pr\left[Z_0 \leq z_{1-\alpha} \mid x = \rho_F + d\right] = \beta \\
&\Leftrightarrow \quad Pr\left[Z + e(\rho_F + d) \leq z_{1-\alpha}\right] = \beta \\
&\Leftrightarrow \quad Pr\left[Z \leq z_{1-\alpha} - e(\rho_F + d)\right] = \beta, \qquad (2.10)
\end{aligned}
$$

where $z_{1-\alpha}$ is the $1-\alpha$ percentile value for the standardized normal distribution; d is a positive value defined as $d \equiv x - \rho_F$, which is called the width of the critical region. The equation (2.10) can be modified so:

$$
\begin{aligned}
-z_{1-\beta} &= z_\beta = z_{1-\alpha} - e(\rho_F + d) \\
\Leftrightarrow -z_{1-\beta} &= z_{1-\alpha} - \frac{d}{\frac{\sigma_\epsilon}{\sqrt{n}}}.
\end{aligned} \tag{2.11}
$$

Further modifying this equation, we get the minimum requirement for the sample size:

$$
n = \left[\frac{(z_{1-\alpha} + z_{1-\beta})\sigma_\epsilon}{d}\right]^2. \tag{2.12}
$$

The meaning of this formula is illustrated in Figure 2.4. For each x, we can draw a probability density function, $f(y)$. The probability of error should be less than α or β when $x \leq SL$ or $x > SL$, respectively; otherwise, the region in which the probability of error exceeds these criteria should be in a critical region with the width of d.

With unknown variance For the unknown variance case, we test the same hypotheses (null and alternative) using another statistic

$$
\begin{aligned}
T_0 &= \frac{U - \rho_F}{\frac{s_\epsilon}{\sqrt{n}}} = \frac{U - x}{\frac{s_\epsilon}{\sqrt{n}}} + \frac{x - \rho_F}{\frac{s_\epsilon}{\sqrt{n}}} = T(n-1) + e'(x) \\
&\approx Z + e'(x),
\end{aligned} \tag{2.13}
$$

where s_ϵ^2 is the sample variance; T follows the Students t-distribution with the degree of freedom at $n - 1$; and $e'(x)$ is a function of x. The approximation in the last line is only reasonable when n is large enough. With the required probabilities of false positive, α, and false negative, β, the required minimum sample size is estimated using the following procedure. If the mean is $x = \rho_F + d$ instead of ρ_F, then the statistic, t_0, has a noncentral t-distribution with a non-centrality

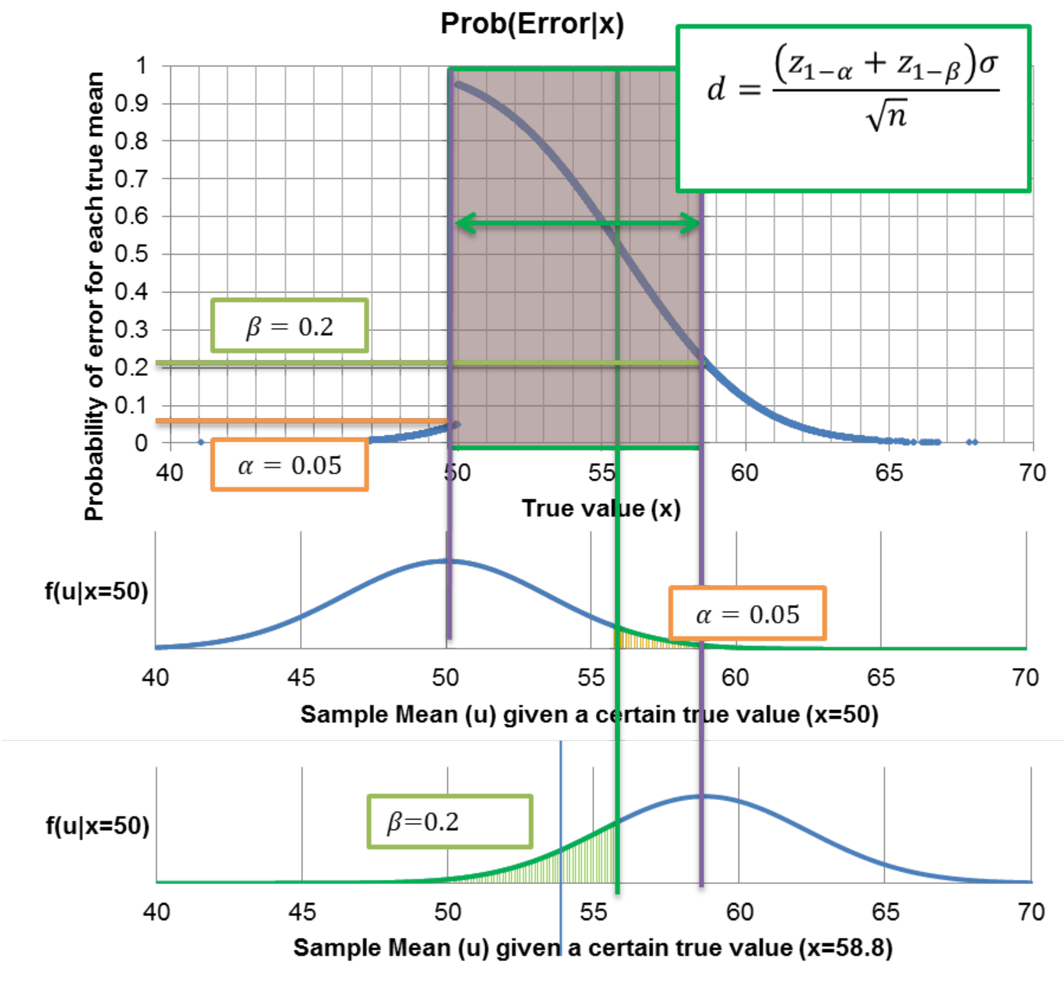

Figure 2.4: Illustration of a critical region with $\alpha = 0.05$ and $\beta = 0.2$

parameter $\Delta = \frac{\rho_F - \rho_F + d}{\frac{\sigma_\epsilon}{\sqrt{n}}} = \frac{d\sqrt{n}}{\sigma_\epsilon}$, where σ_ϵ is not the sample standard deviation, s_ϵ, but rather the population standard deviation for the random variable Y; this variance is still unknown and a nuisance. Similar to the case of a known variance, the minimum requirement for the sample size is

$$n = \left[\frac{(t_{n-1;\,1-\alpha} + t_{n-1;\,1-\beta})\sigma_u}{d} \right]^2, \tag{2.14}$$

where σ_u is a known upper bound for the unknown σ_ϵ. This formula cannot be calculated directly because of the contradictory scenario; the critical value for t-distribution needs a degree of freedom prior to calculating of the minimum sample size. Owen provides tables that identify the required sample size where α, β, and $\Delta = \frac{d}{\sigma_\epsilon}$ are set (Owen, 1962). However, this procedure has two disadvantages: (a) the tables are required and (b) there is a probability that a false positive, α, need to be one of the four values used in the table.

A method used to overcome these disadvantages is Stein's two-stage sampling scheme (Stein, 1945). Desu and Raghavarao (1990) describe the two stage t-test in four steps:

- Obtain an initial sample of size $n_1 (> 2)$. Let s_1^2 be the variance of this sample.
- Let $c = \left(\frac{d}{t_{n_1-1,\alpha} + t_{n_1-1,1-\beta}} \right)^2$.
- Let $n = \max\left(n_1, \lfloor s_1^2/c \rfloor + 1\right)$, where $\lfloor a \rfloor$ is the maximum integer that does not exceed a.
- Take $n - n_1$ additional observations.

Note that we use n and n_1 differently in deriving the critical region of a one-sided α-level test, d, for testing $H_0 : x = \rho_F$ against the alternative, $H_1 : x > \rho_F$, giving a power of statistics of at least $1 - \beta$ at $x = \rho_F + d$ as

$$u > \rho_F + t_{n_1-1,\alpha} \frac{s_1}{\sqrt{n}}, \tag{2.15}$$

where u is the sample mean of sample size, n. Through the above steps, an initial sample size converges to the proper sample size when the initial sample size is smaller than the proper size.

As another method to calculate the minimum sample size, an approximation of the procedure used to estimate the sample size has been suggested. This method, used by the U.S. Environmental Protection Agency (2006), assumes that a new random variable, $W = E(Y) + kS$, is approximately normal with $E(W) = x + k\sigma_\epsilon$ and variance $\sigma_W^2 = \left(\frac{s_\epsilon^2}{n}\right)\left(1 + \frac{k^2}{2}\right)$. The details of deriving the equation were originally described in Eisenhart et al. (1947) and explained well in Guenther (1981). The approximated minimum required sample size is

$$n \geq \frac{[z_{1-\alpha} + z_{1-\beta}]^2 \sigma_\epsilon^2}{d^2} + \frac{1}{2} z_{1-\alpha}^2 \approx \frac{[z_{1-\alpha} + z_{1-\beta}]^2 s_\epsilon^2}{d^2} + \frac{1}{2} z_{1-\alpha}^2, \tag{2.16}$$

where we use the approximation of $\sigma_\epsilon^2 \approx s_\epsilon^2$, which is reasonable when n is large enough. We do not use two sample cases in this thesis. Applications of this normality approximation for a variety of cases are shown in Guenther (1981) and Schouten (1999).

With Required Level of Significance

Consider a situation in which we need to estimate an unknown true value under two safety requirements on type I error and its critical region. This situation can represent SSD for a confidence interval. Let x be the unknown true value. The problem itself is the same as that in the previous subsection; we use the same hypotheses. Similarly, let Y be a random variable that represents the observed sample mean and follow the normal distribution with its mean and variance at x and $\sigma_y^2 = \sigma_z^2/n$, respectively. Note that σ_z^2 is the variance of individual samples. We can set a required level of significance at α; the critical region is $Z < z_\alpha$, where $Z = (Y - x_b)/(\sigma_y)$.

The critical value is derived as

$$
\begin{aligned}
& \frac{y_{cr} - x_b}{\sigma_y} = z_\alpha \\
\Leftrightarrow \quad & y_{cr} = x_b + z_\alpha \sigma_y. \qquad (2.17)
\end{aligned}
$$

If the sample mean is lower than x_{cr}, operators can reject H_0 with α level of significance, and, consequently, they will take a_2, no-action. Otherwise, they should take a_1, replacement.

Remember we have the second safety criteria d, the difference between the population mean, x, and the sample mean, Y. If we assume $x = x_b$ is true, we get the SSD equation as follows:

$$
\begin{aligned}
x_b - y_{cr} \leq d \quad & \Leftrightarrow \quad -z_\alpha \sigma_y \leq d \\
& \Leftrightarrow \quad -z_\alpha \frac{\sigma_z}{\sqrt{n}} \leq d \\
& \Leftrightarrow \quad n \geq \frac{z_\alpha^2 \sigma_z^2}{d^2}. \qquad (2.18)
\end{aligned}
$$

2.4.3 Hypothesis-Testing Approach with Bayesian Probability Updating

The Bayesian approach also offers SSD methods that identify the minimum sample size needed to satisfy given requirements. This approach uses safety criteria that are compared with an average variability of the updated probability density/mass distribution (posterior distribution), such as an average confidence interval of a sample mean. Taking the expectation of values related to a posterior distribution, a decision maker calculates a minimum sample size that meets given requirements for the expected values. The main difference from the frequentist approach is that the Bayesian one uses a prior distribution of an unknown parameter and a likelihood function of data. Comprehensive introductions to these diverse SSD models are provided by Adcock (1997) and Pham-Gia and Turkkan (1992). Later Khalifa et al. (2012) extended this method to a finite population corrosion inspection problem and derived a closed form formula. As another type of combined approach, Xing et al. (2016) proposed a sample size determination method based on Bayesian sequential testing. These approaches still fail to overcome one limitation: they cannot take cost into consideration in a decision-making problem.

Type I error for the mean value, α, and its interval width, l, are used to restrict the acceptable variance of the posterior distribution within three methods: the average coverage criterion (ACC), the average length criterion (ALC), and the worst outcome criterion (WOC). For these methods, l and α are pre-required. With ACC, a covered area of the posterior distribution with l is calculated, and the sample size is determined as the covered area becomes more than $1-\alpha$ percent of all when l covers the distribution properly, which is called the highest posterior density (HPD) interval. With ALC method, $l'(x)$ is derived as the covered area that equals $1-\alpha$ percent of all when $l'(x)$ is HPD, and then the sample size is chosen, as the expectation of $l'(x)$ is less than l. With WOC, the worst case is taken into consideration instead of the expectation about observations. For a normal mean with a known variance case, these three methods (ACC, ALC and WOC) derive the same sample size (Adcock, 1988, 1997). Pham-Gia and Turkkan (1992) suggest three more methods for deriving the minimum sample size needed to satisfy given requirements: 1) they use the posterior variance as a requirement for SSD instead of l and α, 2) they set the maximum acceptable posterior cost, 3) they set the minimum requirement for *EVSI*, although observation cost is not considered in the analysis.

2.5 Comparing Sample Size Determination Methods

Frequentist approach has general versatility; it can be used for parameter estimation with one-sided and two-sided hypothesis testing or confidence interval estimation. However, the method cannot reflect observation cost or prior information about an unknown parameter that we want to estimate. The hypothesis testing method can only answer whether we can determine the hypotheses for the sample size and its obtained data; when a sample mean that is not in a critical region is observed, we can make a decision with a reasonable probability of errors at most α or β. In other words, the method does not give us any instruction if we get a sample mean within the critical region.

Bayesian approaches with such minimum requirements as safety criteria are able to consider prior information and update it with observed data. But these methods are even worse for the two-action problem because they are proposed for two-sided hypothesis testing; thus, we cannot simply apply the methods to the problem. Moreover, these models do not take into account observation costs, so these Bayesian approaches are not suitable for the problem.

The VoI-based SSD method is the way to propose a proper sample size when we want to consider observation costs and prior information about the unknown value. With a cost function, which needs to be proposed additionally, we can derive the optimal sample size based on all the information we have currently.

2.6 Engineering Standards and Guidelines for Sampling Size Selection

For engineering purposes, various types of standards and guidelines for SSD have been provided. Table 2.2 summarizes the approach taken by various standards and guidelines for SSD (American Society for Testing and Materials, 2009, 2015, 2016; Electric Power Research Institute, 1999; U.S. Department of Defence, 1957, 1989; U.S. Environmental Protection Agency, 2006; U.S. Nuclear Regulatory Commission, 1998, 2011). The first four standards are for components in a nuclear power plant, and the other standards and guidelines are used for general components. All can

Table 2.2: Summary of standards and guidelines for sample size determination

Code number	Index	Method	Category
CSA N285.4 (2011)	*	Deterministic	Operational experience
EPRI TR-017218-R1 (1999)	Lot size	Deterministic /Cube-root	Operational experience
U.S. NRC NUREG-1475 (2011)	Percentage of defective products / Mean value	Hypothesis-testing	Frequentist
U.S. NRC NUREG-1505 (1998)	Mean value	Hypothesis-testing	Frequentist
MIL-STD-105D (1989)	Percentage of defective products	Hypothesis-testing	Frequentist
MIL-STD-414 (1957)	Mean value	Hypothesis-testing	Frequentist
ASTM-E122 (2009)	Mean value	Hypothesis-testing	Frequentist
ASTM-F302 (2015)	Lot size	Cube-root	-
U.S. EPA QA/G-4 (2006)	Mean value	Hypothesis-testing	Frequentist

be classified under three categories of methods: hypothesis-testing, deterministic, and cube-root. Hypothesis-testing approaches are based on frequentist statistics and have been used for a long time and in many fields. Deterministic methods rely on operational experiences. The rational behind the suggested sample sizes is not clear. The cube-root method uses an equation; a sample size is the cube-root value of the lot size of components. The theoretical background of this method is not explained by the standards. No standard is based on the VoI concept.

2.6.1 Standards for Nuclear Power Plants in Canada

Guidelines of sampling inspection for nuclear power plants are summarized in CSA N285.4, whose title is "Periodic inspection of CANDU nuclear power plant components" (National Standards of Canada, 2014). It describes sampling at the beginning of use and at each periodic inspection outage. For each component, required sample size and sampling rules are indicated. For example, the minimum sample sizes for baseline (initial) inspection and each inspection interval are set at 15 and 10, respectively. Requirements for other components are summarized in Table 2.3. Although the guidelines have worked well at actual operating sites, no theoretical rationale exists for the sample sizes and requirements for sample selection.

Table 2.3: Sample size requirements for a CANDU reactor

<table>
<tr><th colspan="3" rowspan="2"></th><th rowspan="2">Baseline (Inaugural) inspection</th><th colspan="2">Periodic inspection</th><th rowspan="2">Inspection intervals</th></tr>
<tr><th>Category A</th><th>Category B</th></tr>
<tr><td rowspan="7">General</td><td colspan="2">Piping</td><td>All</td><td>All</td><td>one joint in each pipe</td><td>7< t <12 (about 10-year cycle)</td></tr>
<tr><td colspan="2">Vessels</td><td>All</td><td>All</td><td>>1/3</td><td>7< t <12 (about 10-year cycle)</td></tr>
<tr><td colspan="2">Mechanical couplings</td><td>All</td><td>All</td><td>10%</td><td>7< t <12 (about 10-year cycle)</td></tr>
<tr><td colspan="2">Pumps</td><td>All</td><td>All</td><td>>1/3</td><td>7< t <12 (about 10-year cycle)</td></tr>
<tr><td colspan="2">Valves</td><td>All</td><td>All</td><td>>1/3</td><td>7< t <12 (about 10-year cycle)</td></tr>
<tr><td colspan="2">Supports</td><td>All</td><td>All</td><td>All</td><td>7< t <12 (about 10-year cycle)</td></tr>
<tr><td colspan="2">Rotating machinery</td><td>All</td><td>All</td><td></td><td>7< t <12 (about 10-year cycle)</td></tr>
<tr><td colspan="3">Confirmentary inspection</td><td>>10%</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Corrosion and erosion</td><td colspan="2">Components and systems</td><td></td><td></td><td></td><td>4-year period</td></tr>
<tr><td colspan="2">pumps and valves</td><td></td><td></td><td></td><td>4-year period</td></tr>
<tr><td rowspan="4">Fuel channels</td><td colspan="2">Volumetric and dimensional inspection</td><td>n ≥15</td><td colspan="2">n ≥10</td><td>4< t <8 (about 6-year interval)</td></tr>
<tr><td colspan="2">Hydrogen equivalent concentration (H_eq) determination</td><td></td><td colspan="2">n ≥10 (n ≥6 for interval 1)</td><td>4< t <8 (about 6-year interval)</td></tr>
<tr><td colspan="2">Pressure tube material properties testing</td><td></td><td colspan="2">n ≥1</td><td>4-year interval (after 12-year operation)</td></tr>
<tr><td colspan="2">Material surveillance of fuel channel annulus spacers</td><td></td><td colspan="2">All</td><td>4-year interval (after 12-year operation)</td></tr>
<tr><td rowspan="2">Fuel channel feeder pipes</td><td colspan="2">wall thickness measurement</td><td>20 inlet and 20 outlet</td><td colspan="2">10 inlet and 10 outlet</td><td>6-year interval</td></tr>
<tr><td colspan="2">Feeder pipe visual inspections</td><td>All</td><td colspan="2">1/4 (10 detailed inspection)</td><td>10-year interval</td></tr>
<tr><td rowspan="4">Steam generator (SG) tubes</td><td rowspan="2">Volumetric inspection of tubing</td><td>Steam generators</td><td>25%</td><td colspan="2">n ≥5% or 25 tubes</td><td>6-year interval</td></tr>
<tr><td>Separate preheaters</td><td>25%</td><td colspan="2">n ≥2% or 25 tubes</td><td>10-year interval</td></tr>
<tr><td colspan="2">Secondary-side tube and tube support visual inspections</td><td></td><td colspan="2">One steam generator</td><td>10-year interval</td></tr>
<tr><td colspan="2">Metallurgical examination of tubing</td><td></td><td colspan="2">n ≥1</td><td>6-year interval (after 10-year operation)</td></tr>
</table>

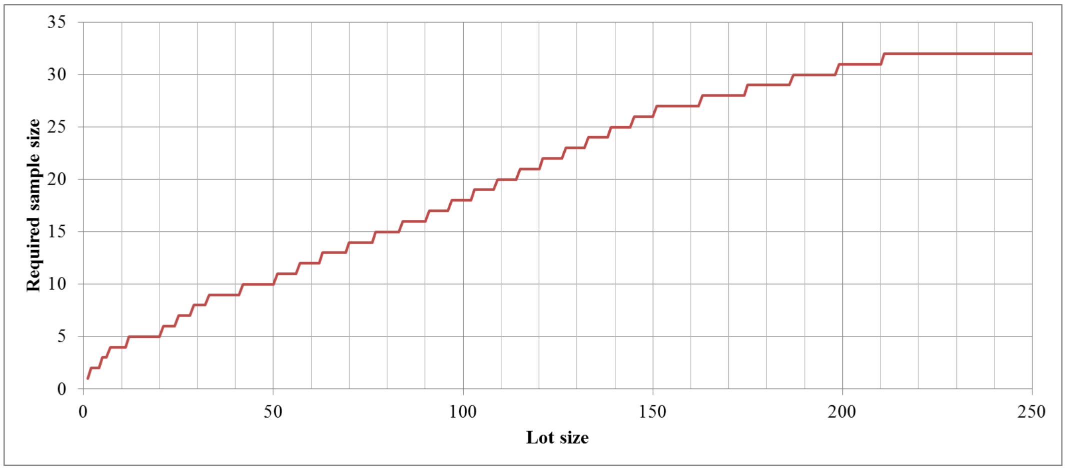

Figure 2.5: Required sample size for each lot size

2.6.2 Standards for Nuclear Power Plants in the United States

Several documents show guidelines of sampling inspection for nuclear power plants in the United States. EPRI TR-017218-R1, "Guideline for sampling in the commercial-grade item acceptance process," specifies a sampling plan and its sample size (Electric Power Research Institute, 1999). For either non-destructive or destructive tests and inspections, sample size is determined by choosing one of three plans, based on the importance of the inspection. The required sample sizes are summarized in tables, for non-destructive or destructive testing. The sample sizes are set referring to the acceptable quality level (limit) (AQL), and to the limiting quality (LQ), which is sometimes called the rejectable quality level (RQL) or lot tolerance percent defective (LTPD), but more-detailed explanations of each number are not included. Figure 2.5 shows how a required sample size increases as the lot size becomes larger. Figure 2.6 depicts an example of how the type I error changes as the lot size increases. Basic analysis of the sample size guideline reveals inconsistencies in its underlying theory.

NUREG-1475 summarizes statistical theory and some methodologies that can be applied to

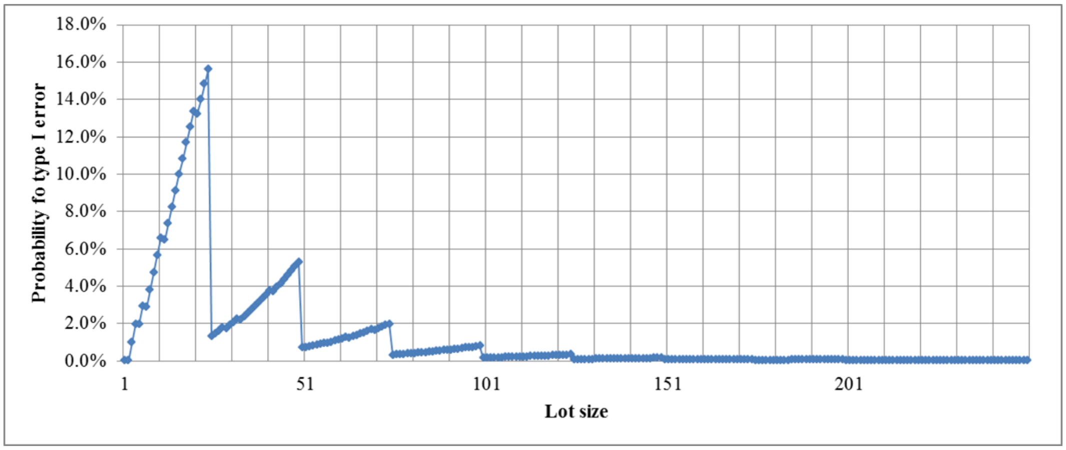

Figure 2.6: Type I error for each lot size

regulating nuclear power plants (U.S. Nuclear Regulatory Commission, 2011). Although this was originally published as a guideline for licensing and regulating nuclear power plants, it is more like a textbook of probability and statistics relating to the area. Both frequentist and Bayesian statistics are described, but only the frequentist SSD approach is detailed. In cases of estimating a mean or comparing two means, a required sample size is derived in a context of hypothesis testing. The standard also derive a SSD approach for testing the percent defective, when the defective occurrence is modelled by a hypergeometric distribution. A required sample size is derived for an example through trial and error, with several combinations of sample sizes and rejection regions of the null hypothesis.

NUREG-1505 is a textbook on using nonparametric statistics for final status decommissioning surveys in nuclear power plants (U.S. Nuclear Regulatory Commission, 1998). This document explains why and in which cases nonparametric statistics is more important than the parametric statistics usually used with assumed normality of obtained data; then it introduces two nonparametric hypothesis testing approaches: the Wilcoxon rank sum test and sign test. How to derive a required sample size is explained for each.

2.6.3 Standards for General Engineering Components in the United States

MIL-STD-105E is titled "Sampling procedures and tables for inspection by attributes" (U.S. Department of Defence, 1989). The guideline was originally used by the military, but has since been widely used in both military and civilian contexts to determine the AQL of products (Juran et al., 1974). The term "quality" represents the percent of defective products or total number of defects per 100 units; the latter can be more than 100 % according to the definition. The process of sampling is modelled using a Bernoulli process, and the probability of acceptance for each possible quality of a product is calculated by using Binomial or Poisson distribution. In the model, AQL needs to be the same as or higher (worse) than the quality at which 95% of products will be accepted. Figure 2.7 provides an example of the operating characteristic curve (OCC), which shows the relationship between the probability of acceptance and the quality of a product. AQL is defined as the quality at which the probability of acceptance is 95% in the OCC. Similarly, the RQL is defined as an unacceptable quality, which is usually compared with the quality with which the probability of acceptance is 10%; in other words, the probability of rejection is 90%. In Figure 2.7, the AQL and RQL are 0.64 and 25.02, respectively. For each combination of AQL and lot size, the required sample size and acceptance/rejection criteria are summarized in the tables; however, the derivation process for the guideline is not fully or clearly explained. As pointed out by Electric Power Research Institute (1999), the required sample sizes increase intermittently; this characteristic may not be fully representative just before or after an increase. This guideline has been adopted in several standards, such as ASTM-B602 (American Society for Testing and Materials, 2016).

MIL-STD-414 specifies sampling plans when a measurement (as opposed to designations of defective or not) is taken and recorded in the sampling procedure (U.S. Department of Defence, 1957). Although the approach is based on the one in MIL-STD-105D, this standard assumes

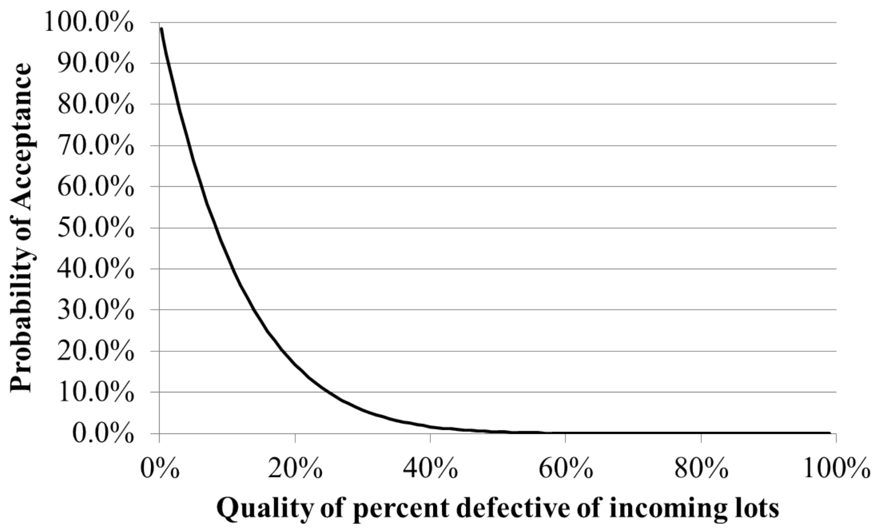

Figure 2.7: Operating characteristic curve of a binomial distribution model with sample of size 8

that measurements follow a normal distribution, and an index calculated from the obtained measurements is compared with a given acceptable value (Juran et al., 1974).

ASTM-E122 provides a required sample size that is based on a consistent statistical background (American Society for Testing and Materials, 2009). The approach, classified as a kind of frequentist approach, is mainly built for estimating a mean value for a certain component's parameter, but it can also be applied to a problem with percentages of defective components. The idea is to find the minimum sample size with which the 3σ range of the sample is the same or narrower than the range of acceptable error, E. Note that the 3σ range covers most probable values of an unknown true mean; it reaches more than 99.7 % of all possible occurrences. The equation for the sample size, n, is as follows:

$$n = \left(\frac{3\sigma_0}{E}\right)^2, \tag{2.19}$$

where σ_0 is an estimate of the lot or process standard deviation of a random variable X.

This has been written for the field sampling of aerospace fluids in containers (American Society for Testing and Materials, 2015). This standard suggests the required sample size for the process, which is calculated with cube-root method. The cube-root method is simply to derive an integer of a cube-root number of a lot size; the results are summarized in Table 2.4. This method has been cited and adapted in Electric Power Research Institute (1999).

Table 2.4: Sample plan introduced in ASTM-F302

Quantity of Containers	Sample
1 to 10	1
11 to 30	3
31 to 70	4
71 to 150	5
151 to 210	6
211 to 530	8
531 to 1170	10

2.7 Gaps in the Research Literature

SSD methods vary, and they should be chosen based on their appropriateness for specific scenarios. Currently, hypothesis-testing as a frequentist approach is used widely; however, the method cannot reflect observation cost or prior information about an unknown parameter that we want to estimate. An SSD method with VoI concept can overcome both weaknesses, so VoI-based methods are appropriate for maintenance problems.

Stochastic degradation models have not applied to maintenance optimization analysis focusing inspection decisions. Most studies assume random variable models or set transition probability among limited numbers of conditions. The studies using VoI concept focus on measurement errors and assume the random variable models. In these studies, decision makers focus on accuracy of periodical inspection or sensor monitoring instead of deciding whether to inspect and/or what sample size to inspect. On the other hand, the studies based on the partially observable Markov decision process (POMDP) assume limited number of conditions, such as no-damage, lightly-damaged, severely-damaged, and failure, and transition probabilities are set for each possible condition-change.

Chapter 3

Diagnostic Inspection of a Component Population

The aim of this chapter is to illustrate a value of information (VoI)-based sample size determination (SSD) approach for diagnostic inspections, with due consideration of the cost consequences of decisions. The proposed approach is formulated within the context of engineering components maintenance. An example is presented, and the parameter sensitivity is illustrated. The proposed approach is compared with the hypothesis-testing approach as well to show the advantages of value of VoI approach.

3.1 Problem Definition

Consider a population of N statistically identical components in an engineering system, which could be vulnerable to some degradation process. A Decision Maker is interested in finding out the extent of degradation in the population and replacing any defective components. Note that the term “defective components” defines components that are going to fail before the next diagnostic inspection. The cost of inspecting a component is C_I. If components experience a degradation failure in service, it would result in a cost, C_F, resulting from both interruption of the operation and repair of the system. During an inspection, a component can be preventively replaced at a cost, C_P, such that $C_P < C_F$.

Because of large population size, the cost of full inspection, $N\,C_I$, is so large that the decision maker prefers not to commit to full inspection at the outset. The reason is that if degradation affects a fairly small number of components, then full inspection would lead to considerable loss of inspection resources (costs and time). A preferred approach is to demonstrate via sampling that the extent of degradation small enough that full inspection is not warranted.

Thus, the decision maker decides to inspect a small sample, n, $n < N$, and use the information obtained to make a decision to take one of the following two actions:

a_1 : Inspect remaining population of, $(N - n)$, and replace all defective components (Full inspection)

a_2 : Do not inspect remaining component population, and let components fail in service incurring a cost, C_F per failure (Do-nothing option)

A key objective is to find an optimal sample size, n, to support this decision problem.

3.1.1 Percentage of Defective Components

This section describes how to derive the percentage of defective components as the representative random variable without explicit consideration of time in the model. Assume that the degradation level at time t, $Y(t)$, follows a random rate model:

$$Y(t) = Rt \tag{3.1}$$

where R is the random corrosion rate that reflects the variability observed in a sample of deterioration data in a population of similar components. Consider the random variable, R, which follows a probability density function of $f(r)$ and has a cumulative function of $F(r)$. A component will fail before the next diagnostic inspection at t if it has a higher corrosion rate than a threshold ρ, with which a component will fail at t. The probability of failure is calculated as

$$P_f = 1 - F(\rho). \tag{3.2}$$

This probability can also be interpreted as the percentage of defective components in the population. When the function, $F(r)$, is not obvious, the rate, P_f, needs to be treated as a random variable, and can be denoted such as X. Thus, the percentage of defective components can represent the variable nature of the deterioration in a population of similar components.

3.1.2 Cost Functions

The cost associated with the two actions depends on the percentage of degraded or defective components, X, in the population. Unless the full inspection of N components is performed, the defective fraction remains unknown to the decision maker, and so it is treated as a random variable. The percentage of defective components can be the representative random variable of this problem.

The two cost functions corresponding to the two actions for a certain value, $X = x$, are defined as follows:

$$C(a, x \mid n, w) = \begin{cases} C(a_1, x \mid n, w) \\ C(a_2, x \mid n, w) \end{cases} = \begin{cases} (N-n)C_I + [(N-n)x + w]C_P \\ (N-n)xC_F + wC_P, \end{cases} \tag{3.3}$$

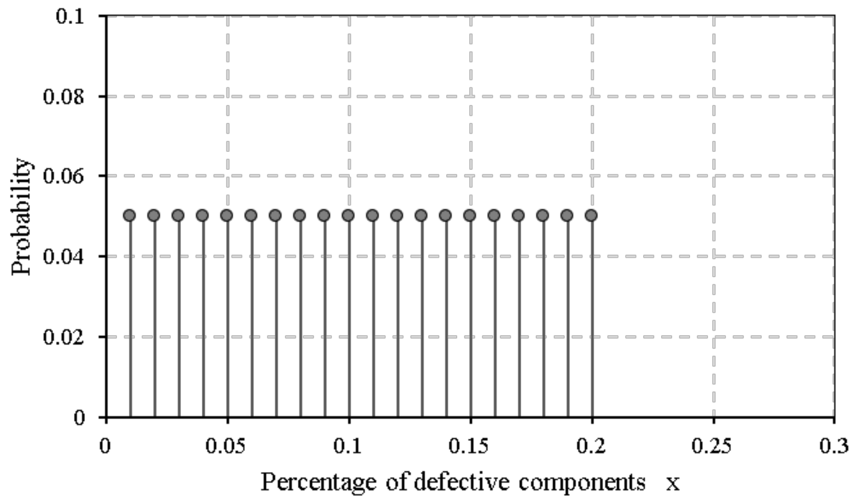

Figure 3.1: Prior distribution for percentage of defective components.

where w is the number of detected defective components in the inspection sample of size n. Note that these cost functions do not include the cost of sampling, nC_I, as it is a deterministic cost.

In the Bayesian framework, a prior distribution can be assigned to X, which can be updated as information become available. A discrete probability mass function is chosen to model X, as shown in Figure 3.1, and defined below:

$$f_X(x_i) = \begin{cases} 0.05 & \text{for } x_i = i/100,\ i = 1, 2, \cdots, 20 \\ 0 & \text{Otherwise.} \end{cases} \tag{3.4}$$

This distribution is chosen for illustration and it is also used later in the numerical example. The mean and standard deviation of this distribution are 0.105 and 0.0577, respectively.

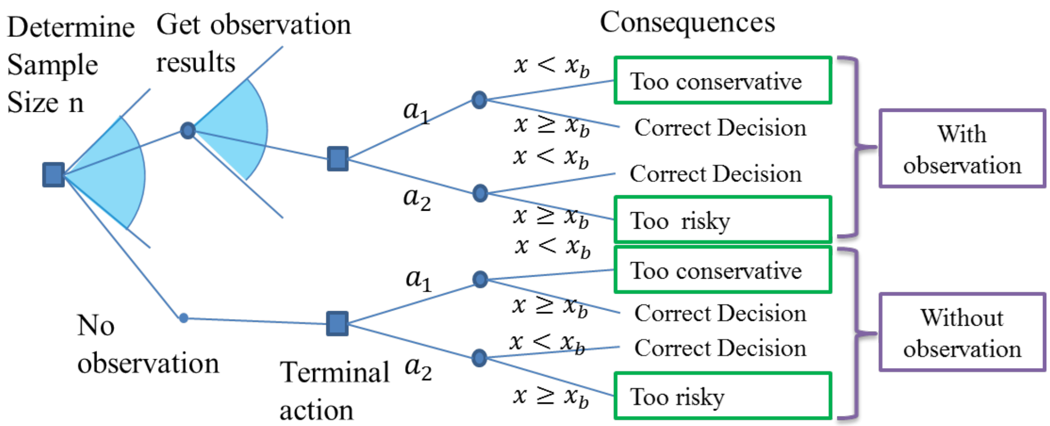

Figure 3.2: Extended decision tree for a maintenance problem as a Bayesian pre-posterior analysis.

3.2 Value of Information Analysis

The basis for formulating the VoI problem is the Bayesian pre-posterior analysis. The term "Pre-posterior" is used to denote the consideration of a posterior situation, while the decision maker is actually still in a prior situation. A component replacement problem is illustrated as Figure 3.2. This analysis determines the options of n and a that minimize expected costs on unknown values w and x.

3.2.1 Prior Analysis: No Inspection Data

In the absence of any inspection sample data, the decision maker should choose that action which minimizes the expected cost. The prior cost function are defined as follows:

$$C(a,x) \equiv \begin{cases} C(a_1, x \mid n=0, w=0) \\ C(a_2, x \mid n=0, w=0) \end{cases} = \begin{cases} NC_I + NxC_P \\ NxC_F. \end{cases} \tag{3.5}$$

The variation of these costs with x is shown in Figure 3.3. In this two-action problem, there is a break-even value, $X = x_b$, for which the two costs are equal, and it is calculated as

$$\begin{aligned}(N-n)C_I + (N-n)x_b C_P &= (N-n)x_b C_F \\ \Leftrightarrow x_b &= \frac{C_I}{C_F - C_P}.\end{aligned} \tag{3.6}$$

In the VoI approach, the break-even value, x_b, is a key parameter for decision-making. If $x > x_b$, the optimal action is a_1, full-inspection. In case of $x < x_b$, the optimal action is a_2, "do-nothing". The expected prior cost for each terminal action is calculated as follows:

$$\mathbb{E}_X\left[C(a,X)\right] = \begin{cases} \mathbb{E}_X\left[C(a_1,X)\right] \\ \mathbb{E}_X\left[C(a_2,X)\right] \end{cases} = \begin{cases} N x_b (C_F - C_P) + N\bar{x}C_P \\ N\bar{x}C_F, \end{cases} \tag{3.7}$$

where $\bar{x}$ is the mean of X. For modifying equations to simpler form, we use $C_I = x_b(C_F - C_P)$, which is originally derived in Equation (3.6). The optimal action should lead to a smaller cost than the other option; thus, if a_1 is optimal, the necessary condition is derived as

$$\begin{aligned}N x_b (C_F - C_P) + N\bar{x}C_P &< N\bar{x}C_F \\ \Leftrightarrow \bar{x} &> x_b.\end{aligned} \tag{3.8}$$

If the mean value of the prior distribution is more than the break-even value, x_b, the optimal action is a_1, full-inspection. The minimum cost with the prior distribution is calculated thus:

$$\min_a \mathbb{E}_X\left[C(a,X)\right] = \begin{cases} N x_b (C_F - C_P) + N\bar{x}C_P & \text{if } \bar{x} > x_b \\ N\bar{x}C_F & \text{if } \bar{x} \le x_b, \end{cases} \tag{3.9}$$

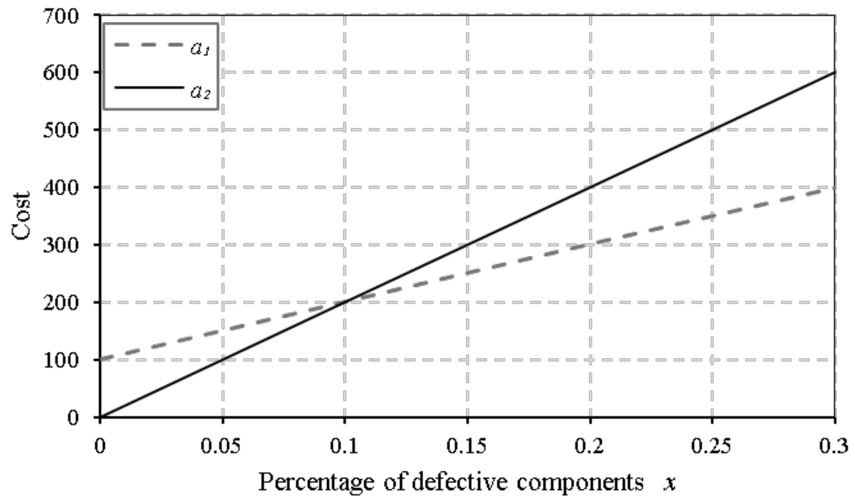

Figure 3.3: The prior cost functions for the two actions, as a function of defective fraction.

3.2.2 Posterior Analysis

After observing a sample, the decision maker derives a posterior distribution of X. Assuming that the number of defective components, r, in a sample of n, follows a binomial distribution, its mass function can be written as

$$\begin{aligned} f_W(w \mid n) &= \sum_{x \in X} f_{W|X}(w \mid n, x_j) f_X(x_j) \\ &= \sum_{j=1}^{20} \binom{n}{w} x_j^w (1 - x_j)^{n-w} \cdot f_X(x_j). \end{aligned} \tag{3.10}$$

The posterior distribution for X is calculated using Bayes' rule. For a value of w $(0 \leq w \leq n)$ and x_k, the updated distribution is given as:

$$\begin{aligned} f_{X|W}(x_k \mid n, W) &= \frac{f_{W|X}(w \mid n, x_k) f_X(x_k)}{f_W(w \mid n)} \\ &= \frac{x_k^w (1-x_k)^{n-w} \cdot f_X(x_k)}{\sum_{j=1}^{20} x_j^w (1-x_j)^{n-w} \cdot f_X(x_j)}. \end{aligned} \tag{3.11}$$

As an example, posterior distributions of X are computed for $n = 6$ and $w = 0, 2, 4, 6$, and plotted in Figure 3.4. The expected posterior cost is calculated as

$$\begin{aligned} \mathbb{E}_{X|W}\left[C(a, X \mid n, w)\right] &= \begin{cases} \mathbb{E}_{X|W}\left[C(a_1, X \mid n, w)\right] \\ \mathbb{E}_{X|W}\left[C(a_2, X \mid n, w)\right] \end{cases} \\ &= \begin{cases} (N-n)x_b(C_F - C_P) + (N-n)\bar{x}''(n,w)C_P + wC_P \\ (N-n)\bar{x}''(n,w)C_F + wC_P, \end{cases} \end{aligned} \tag{3.12}$$

where $\bar{x}''(n, w)$ indicates a mean value of the probability mass function of x posterior to knowing $W = w$ and is a function of n and w. Note that the posterior distribution is still a discrete distribution. Similar to Equation (3.8), when $\bar{x}''(n, w) < x_b$, we take a_1 as the optimal action; otherwise, we take a_2. For a given n, if we set h_n, which satisfies $\bar{x}''(n, h_n) < x_b$ and $\bar{x}''(n, h_n+1) > x_b$ simultaneously, we can calculate the minimum expected posterior cost so:

$$\begin{aligned} &\min_a \mathbb{E}_{X|W}\left[C(a, X \mid n, w)\right] \\ &= \begin{cases} (N-n)x_b(C_F - C_P) + (N-n)\bar{x}''(n,w)C_P + wC_P & \text{if} \quad w > h_n \\ (N-n)\bar{x}''(n,w)C_F + wC_P & \text{if} \quad w < h_n. \end{cases} \end{aligned} \tag{3.13}$$

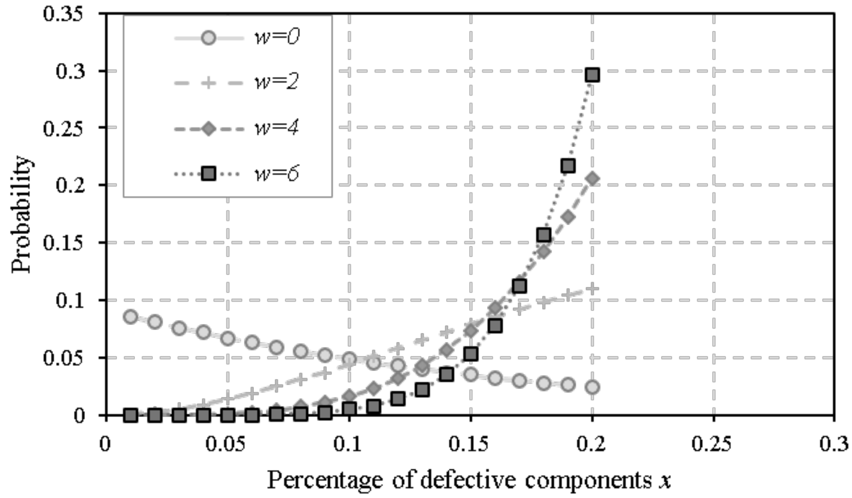

Figure 3.4: Posterior distributions of the percentage of defective components: An illustration for $w = 0, 2, 4$, and 6.

3.2.3 Pre-posterior Analysis

The next step is the pre-posterior analysis, in which the decision maker considers all possible outcomes of sampling in the following manner:

$$
\begin{aligned}
& \mathbb{E}_W \left[\min_a \left\{ \mathbb{E}_{X|W} \left[C(a, X \mid n, w) \right] \right\} \right] \\
= & \sum_{w=0}^{h_n} \left[(N-n)\bar{x}''(n,w) C_F + w C_P \right] \cdot f_W(w \mid n) \\
& + \sum_{w=h_n+1}^{n} \left[(N-n) x_b (C_F - C_P) + (N-n)\bar{x}''(n,w) C_P + w C_P \right] \cdot f_W(w \mid n) \qquad (3.14)
\end{aligned}
$$

EVSI is a positive comparison between prior and posterior expected costs; this criterion represents the value of sampling-inspection before the terminal decision. A more-detailed definition is as follows:

$$\begin{aligned} EVSI(n) &\equiv \mathbb{E}_X\left[C(a^o, X)\right] - \mathbb{E}_W\left[\mathbb{E}_{X|W}\left[C(a^o(w), X \mid n, w)\right]\right] \\ &= \min_a\{\mathbb{E}_X\left[C(a, X)\right]\} - \mathbb{E}_W\left[\mathbb{E}_{X|W}\left[\min_a C(a, X \mid n, w)\right]\right] \end{aligned} \tag{3.15}$$

where $\mathbb{E}_{X|Z}[g(X)]$ is an expectation of $g(X)$ on X conditional on a given condition of $Z = z$; $f_{X|W}(x \mid n, w)$ is a posterior probability density or mass function of a random variable, X, conditional on n and w; a^o is the optimal action without observation; and $a^o(w)$ is the optimal action under a known sampling-inspection result, w. Since *EVSI* is an expected value prior to additional observations, the second term is an expectation for both an uncertain event and the as-yet-unknown observations.

Thus, with Equation (3.15), the *EVSI* is derived as follows:

$$\begin{aligned} & EVSI(n) \\ = \; & \min_a[\mathbb{E}_X\left[C(a, X)\right]] - \mathbb{E}_W\left[\min_a\left\{\mathbb{E}_{X|W}\left[C(a, X)\right]\right\}\right] \\ = \; & \begin{cases} \left[N - \sum_{w=h_n+1}^{n}(N-n)f_W(w \mid n)\right] x_b(C_F - C_P) & \\ \quad - \sum_{w=0}^{h_n}\left[(N-n)\bar{x}''(n,w)(C_F - C_P)\right] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\ - \sum_{w=h_n+1}^{n}(N-n)x_b(C_F - C_P)f_W(w \mid n) + n\bar{x}(C_F - C_P) & \\ \quad + \sum_{w=h_n+1}^{n}\left[(N-n)\bar{x}''(n,w)(C_F - C_P)\right] \cdot f_W(w \mid n) & \text{if } \bar{x} \le x_b. \end{cases} \end{aligned} \tag{3.16}$$

EVSI is not a sufficient objective for SSD; the net gain through the observations should be optimized. Subtracting the costs relating to a sampling-inspection, $Cost(n, w)$, from *EVSI*, we define another net benefit as the *ENGS*:

$$ENGS(n) = EVSI(n) - Cost(n, w) \tag{3.17}$$

This *ENGS* is a function of sample size n, and we estimate the optimal sample size that maximizes the *ENGS*. With Equation (3.17), *ENGS* is derived as

$$\begin{aligned} & ENGS(n) \\ = \; & EVSI(n) - nC_I \\ = \; & \begin{cases} (N-n)C_I \sum_{w=0}^{h_n} \left[1 - \frac{\bar{x}''(n,w)}{x_b}\right] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\ (N-n)C_I \sum_{w=h_n+1}^{n} \left[\frac{\bar{x}''(n,w)}{x_b} - 1\right] \cdot f_W(w \mid n) - nC_I \left(1 - \frac{\bar{x}}{x_b}\right) & \text{if } \bar{x} \le x_b. \end{cases} \end{aligned} \tag{3.18}$$

Note that the details of analytical derivations for *EVSI* and *ENGS* are presented in the Appendix A.

3.2.4 Numerical Example

To demonstrate the proposed VoI approach, the following parameter values are chosen: $N = 100$, $C_F = 20$, $C_P = 10$, and $C_I = 1$. The prior distribution of defective fraction, X, is shown in Figure 3.1. The prior values of expected costs of the two actions, a_1 (full inspection) and a_2 (no inspection) are estimated as 205 and 210, respectively. Thus, a_1, i.e., full inspection of the component population followed by the preventive replacement of defective components found during the inspection is the best prior action.

The next step is the pre-posterior analysis, which is illustrated for a sample size of $n = 6$. Using the procedure described in the previous section, expected of best best action based on sampling results is estimated as,

$$\mathbb{E}_W \left[\min_a \left\{\mathbb{E}_{X|W}\left[C(a_1, X \mid n = 6, w)\right]\right\}\right] = 190.53.$$

Now, the *EVSI* and *ENGS* for $n = 6$ are computed as,

$$\begin{aligned} EVSI(n=6) &= \mathbb{E}_{X|W}\left[C(a_1, X \mid n=6, w)\right] - \mathbb{E}_W\left[\mathbb{E}_{X|W}\left[C(a^o(w), X \mid n=6, w)\right]\right] \\ &= 14.47, \\ ENGS(n=6) &= 8.47. \end{aligned}$$

In this manner, the $ENGS(n)$ is computed for $n = 0$ to $n = 100$ and plotted in Figure 3.5. The optimal sample siz is $n = 20$ for which the *ENGS* takes a maximum value of 11.84. Note that the small mounds in each line occur because the inspection outcome is discrete; once the observed ratio, w/n, overlaps the break-even point, x_b, the *ENGS* line starts to form another small mound.

In a practical setting this result can be used in the following manner. The decision maker inspects a random sample of 20 components. Depending on the number of defectives (r), the following two actions are available:

- If $w \geq 2$, i.e., $w/n \geq x_b(= 0.1)$, then choose a_1, which means full-inspection of the remaining population of 80 components;
- If $w \leq 1$, i.e., $w/n \leq x_b(= 0.1)$, then choose a_2, which no additional inspection required.

3.3 Sensitivity Analysis

This Section evaluates the sensitivity of the optimal sample size to parameters like the break-even value, x_b, the prior mean, $\bar{x}'$, the inspection cost, C_I, and the prior standard deviation, $s(x)'$.

Let us set a base-line situation in which $x_b = 0.1$, $\bar{x}' = 0.105$, $C_I = 1$, and $s(x)' = 0.0577$. Each parameter is shifted from 80% to 120% of its base-line value; all other parameters are fixed.

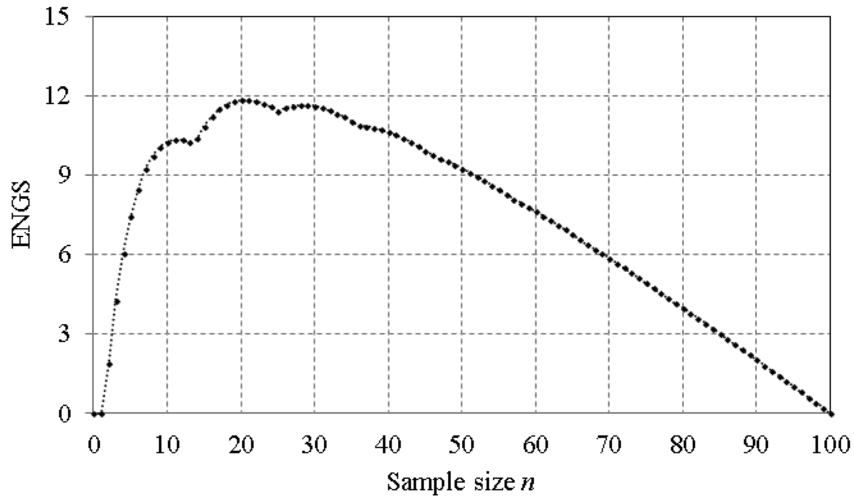

Figure 3.5: A plot of *ENGS* versus the sample size

Figure 3.6 illustrates the fluctuation of optimal sample sizes with changes in parameter values. Note that the two cases, 80% for $\bar{x}'$ and 120% for $s(x)'$, remain blank since a prior distribution becomes improper in these cases; the distribution of x, which must be within the range between 0 and 1, has a positive probability mass for $x < 0$.

3.3.1 Break-Even Value, x_b

The optimal sample size, n^o, for each x_b is calculated and shown in Figure 3.7. The optimal sample size tends to increase as x_b decreases. When x_b becomes smaller, the decision criterion, h_n, decreases since only small numbers of defective components can make the posterior mean, $\bar{x}''(n, w)$, be lower than x_b. Thus, decision makers have more probability of observing w that is higher than the decision criterion, h_n, and consequently choosing a_1 as a terminal action (see Equation (3.13)). The higher the probability of choosing a_1 as a terminal action after sampling inspection, the more the incentive for decision makers to reduce the size of sampling inspection.

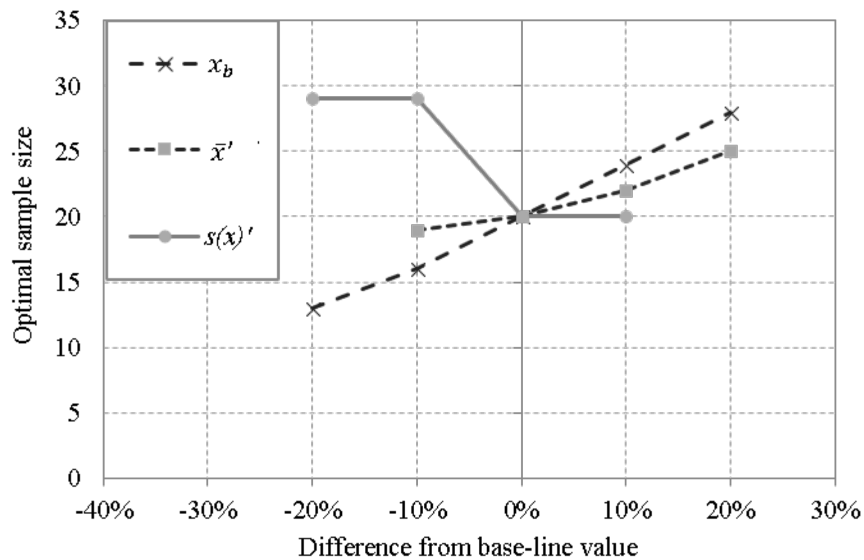

Figure 3.6: Overall sensitivity analysis for parameters

An inverse movement for the optimal sample size exists around $x_b = 0.038$. This exceptional move occurs because of small mounds on the *ENGS* plots (see Figure 3.5). When $x_b < 0.038$, the peak is in the second mound, and the peak moves to the first mound when $x_b = 0.038$. If the criterion, x_b, is greater than 0.136, the best behaviour for the decision maker is to take a_2, no-action, without any sample, which is a sampling inspection.

Figure 3.8 shows $ENGS(n^o)$ for each x_b. The peak of $ENGS(n^o)$ value is obtained when x_b is the same as the prior mean, $\bar{x}' = 0.105$. Mathematically, this is dome because the derivative of *ENGS* with respect to x_b is positive when $x_b < \bar{x}'$ and is negative when $x_b > \bar{x}'$ (see Equation (3.18)).

As x_b is set close to $\bar{x}'$ but is lower than $\bar{x}''$, decision makers have more chance to obtain a smaller $\bar{x}''$ than x_b, which indicates that the terminal action should be different from the action chosen based on prior analysis. In other words, in this case, the impact of additional information is relatively greater than it would be in other situations.

3.3.2 Mean of Prior Distribution, $\bar{x}'$

With fixed values of C_I, N, x_b, and $s(x)'$ at 1, 100, 0.10, and 0.0432, respectively, the influence of the probability mean, $\bar{x}'$, is analysed in Figures 3.9 and 3.10, which are closely related. The results of sensitivity analysis on x_b and $\bar{x}'$ are roughly symmetrical. Similar to the sensitivity analysis on x_b, the highest $ENGS(n^o)$ is obtained when $\bar{x}'$ equals the break-even value, $x_b = 0.10$. *ENGS*s decrease as $\bar{x}'$ deviates from $x_b = 0.10$. The gap between these two parameters represents the importance of additional information, although this is not directly shown in Equation (3.18); when the gap equals zero, the *ENGS* is maximized. If the optimal action is obvious only with prior information, where a decision maker has a low probability of obtaining inspection outcomes that suggest the decision maker change the terminal action, the decision maker need not obtain any sample.

3.3.3 Inspection Cost, C_I

As shown in Equation (3.18), the *ENGS* is proportional to C_I, and the optimal sample sizes of all cases are the same, $n^o = 20$. Thus, in this stated problem, we can normalize *ENGS* by dividing it by C_I or simply set C_I as 1. Decision makers need to consider not the abstract values of C_I, C_F, and C_P but a ratio between C_I and $(C_F - C_P)$; x_b is a key parameter in the stated problem.

3.3.4 Summary

Several insights have been found in this section. x_b is the most sensitive parameter. Decision makers need to identify relationships among C_F, C_P, and C_I, which derive x_b. The relationship

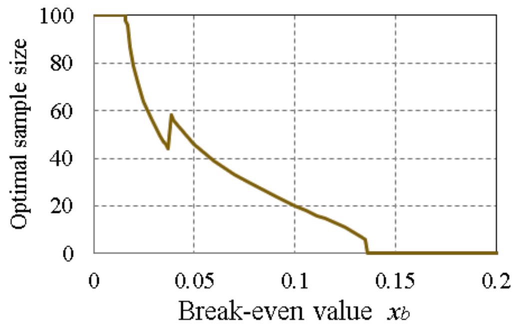

Figure 3.7: Optimal sample sizes vs. x_b

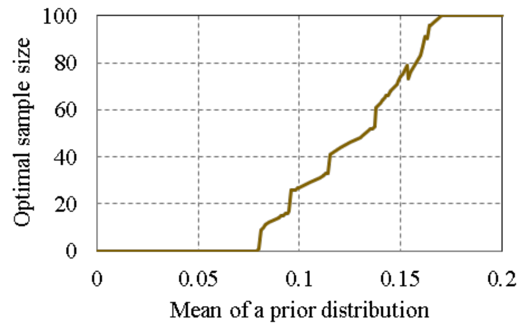

Figure 3.9: Optimal sample size vs. $\bar{x}'$

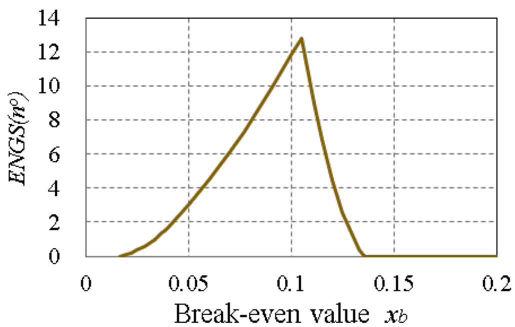

Figure 3.8: $ENGS(n^o)$ vs. x_b

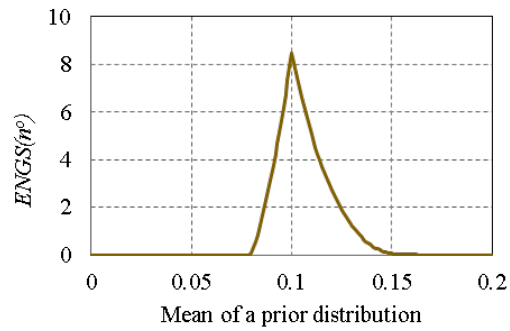

Figure 3.10: $ENGS(n^o)$ vs. $\bar{x}'$

between x_b and $\bar{x}'$ sets the width of n^o values, in which the additional sampling is meaningful. The more vague and difficult terminal decision-making with only prior information is, the more $ENGS(n^o)$ a decision maker obtains; in that situation, additional information has more value for a decision maker. In other words, if the prior information is enough to decide an action to take, a decision maker has no incentive to obtain any additional sample.

3.4 Hypothesis-Testing Approach

Classical hypothesis-testing, a frequentist approach, has primarily been used to estimate the sample size required to test a statistical hypothesis regarding the magnitude of defective fraction, X. In problems related to environmental risk assessment, the U.S. Environmental Protection

Agency (2006) recommends the hypothesis-testing approach. Here, the sample size is selected to ensure that statistical errors of Type I and II stay within certain reasonable limits.

There is a smaller range of X in which the two requirements are not satisfied, which called a gray region. With the gray region's maximum acceptable width as the third criteria, we can determine the minimum sample size that satisfies the three restrictions.

3.4.1 Sample Size Analysis

The stated two-action problem can be translated as a hypothesis-testing problem. Respectively, the null hypothesis and alternative hypothesis are

$$\begin{aligned} H_0: &\quad x = x_b \text{ (or } x \geq x_b), \quad \text{(then take } a_1\text{, full-inspection)} \\ H_1: &\quad x = x_1 = x_b - d \text{ (or } x < x_b), \quad \text{(then take } a_2\text{, no-action)}, \end{aligned}$$

where d is the gap between the two hypotheses, H_0 and H_1, and is defined as $d = x_b - x_1$. The decision for the hypothesis-testing here is followed by an option; if we reject the null hypothesis, the decision maker will take a_2, "No-action." If we fail to reject the null hypothesis, the decision maker should take a_1, "Full-inspection." We can set a safety criteria, α, for the maximum acceptable probability of rejecting H_0 in error when H_0 is true, which is called "type I error." It is assumed here that the normal distribution is an adequate approximation of the sampling distribution of the defective fraction.

In the hypothesis-testing approach, a critical value (x_{cr}), which is the cut-off value for type I error, works as x_b in the VoI-based approach, and is derived as

$$x_{cr} = x_b - z_{1-\alpha}\sqrt{\frac{x_b(1-x_b)}{n}\frac{N-n}{N-1}}. \tag{3.19}$$

If the sampling inspection result, r/n, is lower than x_{cr}, decision makers can reject H_0 with a less than α probability of type I error, and, consequently, they will take a_2, no-action. Otherwise, they should take a_1, full inspection. Note that x_{cr} is always lower than x_b. Similarly, we can set β as the maximum acceptable probability, erroneously not rejecting H_0 when $x = x_b - d$, which means H_1 is true (a type II error). In terms of the power of statistics, the requirement can be explained as the required power of statistics, $1 - \beta$, at $x = x_b - d$. Here, d can be translated as the width of a gray region; in it ($x_b - d \leq x < x_b$), our decisions have a higher probability of errors than β; decision makers have the probability of making type II errors (full-inspection when x is safer, $x < x_b$) more than β when they obtain the outcome of a sampling inspection that is $r/n > x_{cr}$ while the true value, x, is in a gray region ($x_b - d \leq x < x_b$). The relationship among α, β, and d, is illustrated in Figure 3.11. For example, when $\alpha = 0.05$, $\beta = 0.2$, and $d = 0.04$, the cases with $n = 20$, $n = 40$, and $n = 60$ do not satisfy the requirements, whereas the other cases do meet them. Details of how to decide on an action based on an obtained sample is explained by Higo and Pandey (2016), although they use normal distribution instead of binomial distribution.

Through derivation processes explained by the U.S. Environmental Protection Agency (2006), assuming X follows the normal distribution with its mean and variance at x and $x(1-x)(N-n)/(N-1)$, respectively, we get the minimum requirement for the sample size:

$$\begin{aligned} n_{min} \cdot \frac{N-1}{N-n_{min}} &= \left[\frac{z_{1-\alpha}\sqrt{x_b(1-x_b)} + z_{1-\beta}\sqrt{(x_b-d)(1-x_b+d)}}{d}\right]^2 \\ \Leftrightarrow n_{min} &= \frac{N\left[\frac{z_{1-\alpha}\sqrt{x_b(1-x_b)}+z_{1-\beta}\sqrt{(x_b-d)(1-x_b+d)}}{d}\right]^2}{N-1+\left[\frac{z_{1-\alpha}\sqrt{x_b(1-x_b)}+z_{1-\beta}\sqrt{(x_b-d)(1-x_b+d)}}{d}\right]^2}, \end{aligned} \tag{3.20}$$

where $\sqrt{(N-n)/(N-1)}$ is the finite population correction factor, which improves the accuracy in the binomial approximation for a finite population (Sandiford, 1960).

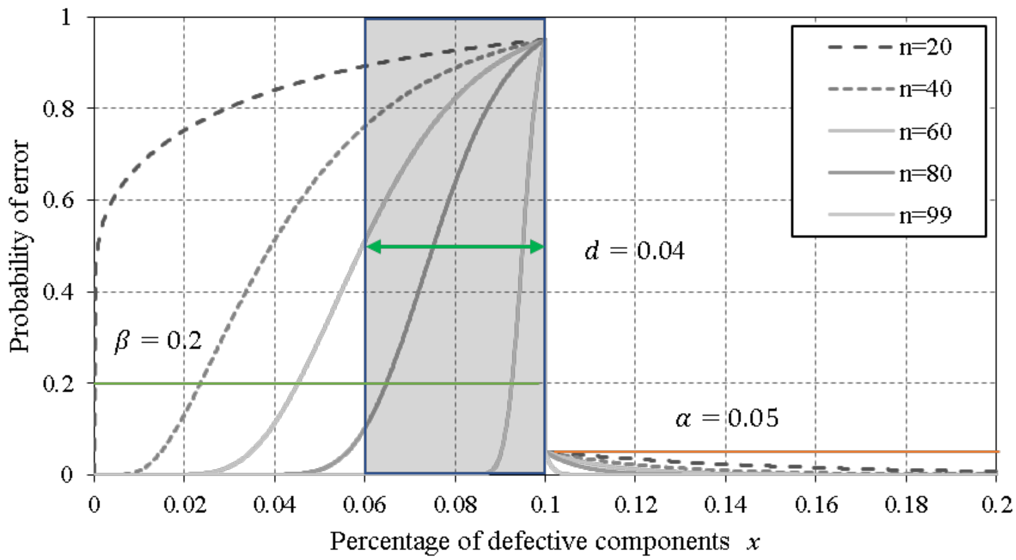

Figure 3.11: Illustration of three safety criteria and probability of error for each sample size ($\alpha = 0.05$, $\beta = 0.2$, and $d = 0.04$).

3.4.2 Numerical Example

The hypothesis-testing approach employs a widely used criteria, called the "20/80 rule", which implies that a sample estimate is within 20% of an unknown population parameter with 95% confidence and 80% statistical power. We can use this rule with reference to the break-even value, x_b, of the defective fraction. To apply this rule, use $\alpha = 0.05$, $\beta = 0.20$, $d = 0.02$, and $x_b = 0.1$ in Equation 3.20, which leads to a required sample size $n \geq 92.9$. Thus, the minimum sample size is 93.

The calculated sample size is a function of the width of the grey region, d, as shown in Table 3.1,where d is varied from 10 to 90 percent of the break-even value of $x_b = 0.10$. As the gray region becomes narrower, the required sample size increases and approaches the lot size, 100.

Table 3.1: Minimum sample size for each d (for $\alpha = 0.05$ and $\beta = 0.20$)

d	$d/x_b(\%)$	n_{min}	x_{cr}
0.09	90	30	0.024
0.08	80	38	0.036
0.07	70	46	0.046
0.06	60	55	0.055
0.05	50	65	0.064
0.04	40	76	0.072
0.03	30	85	0.079
0.02	20	93	0.086
0.01	10	99	0.093

3.5 Comparison

How to choose a sample size and an action at each decision node and react to any possible inspection outcome can be summarized as a policy. The policy contains a selected sample size, and a decision criterion, which is x_b for the VoI approach and x_{cr} for the hypothesis-testing approach. The policies obtained through the two SSD approaches are compared in two scenarios that are simulated based on x, which is either deterministic or stochastic. The deterministic case highlights how a decision maker's bias for x affects the consequences of different policies. Through the comparison, effective situations for each method can be obtained and summarized. The stochastic case provides an overall comparison of the two approaches. When the prior distribution is reasonable, the expected total cost represents how much better one SSD method is than the other.

3.5.1 Evaluation Scheme

In the deterministic case, the true proportion of the defective components, x, is assumed to be a given fixed value. After an inspection policy is obtained through each approach "without" knowing the exact value of x (with only the prior distribution), the policy is applied to the scenario that is calculated "with" the given true value, x. The consequence of each policy is evaluated as a total cost and the two policies are compared. For the VoI approach, given n^o, x_b, and x, the total cost is calculated as follows:

$$\begin{aligned} C_{VoI} &= n^o C_I + \mathbb{E}_{W|X}\left[\min_a C(a, x \mid n^o, w)\right] \\ &= n^o C_I + \sum_{w=0}^{h_{n^o}} C(a_2, x \mid n^o, w) f_{W|X}(w \mid x, n^o) \\ &\quad + \sum_{w=h_{n^o}+1}^{n^o} C(a_1, x \mid n^o, w) f_{W|X}(w \mid x, n^o). \end{aligned} \tag{3.21}$$

Similarly, for the hypothesis-testing approach, given n_{min}, x_{cr}, and x, the total cost is derived as follows:

$$\begin{aligned} C_{HT} &= n_{min} C_I + \sum_{w=0}^{k} C(a_2, x \mid n_{min}, w) f_W(w \mid x, n_{min}) \\ &\quad + \sum_{w=k+1}^{n_{min}} C(a_1, x \mid n_{min}, w) f_{W|X}(w \mid x, n_{min}), \end{aligned} \tag{3.22}$$

where k is the decision criterion, defined as $k = \lfloor n_{min} \cdot x_{cr} \rfloor$. Note that $\lfloor x \rfloor$ is the floor function, which returns the closest integer of less than x.

The stochastic case represents the situation where all information about x can be summarized as a prior distribution (a probability mass function of x). In this case, given the probability of occurrence for each possible x, the expected total cost is derived; for example, that for the VoI approach is

$$\begin{aligned}\mathbb{E}_X\left[C_{VoI}\right] &= n^o C_I + \sum_{x \in X} \sum_{w=0}^{h_{n^o}} C(a_2, x \mid n^o, w) f_{W|X}(w \mid x, n^o) f_X(x) \\ &\quad + \sum_{x \in X} \sum_{w=h_{n^o}+1}^{n^o} C(a_1, x \mid n^o, w) f_{W|X}(w \mid x, n^o) f_X(x). \end{aligned} \tag{3.23}$$

3.5.2 Certain x Case

For each possible x, the costs for each approach are calculated and summarized in Figure 3.12. The VoI approach reduces the cost if x is less than x_b, which is set at 0.10. However, once x becomes larger than x_b, the VoI approach raises the cost, since the decision can be wrong if the sample of a sampling-inspection does not represent the group. This result indicates that if the decision maker sets a prior distribution conservatively, the VoI approach can provide a lower total cost. For instance, if the true percentage of defective components is 0.01, and we set a prior distribution with its mean at 0.105, the difference between the cost with VoI approach is less than the cost with the hypothesis-testing approach with the $d = 0.02$ case by more than 18 times C_I.

3.5.3 Uncertain x Case

With the prior distribution shown in Figure 3.1 as an example of reasonable estimation for the occurrence of each $X = x$, the expected costs for each approach are compared (Table 3.2). The results indicate that the VoI approach is best when the prior distribution is proper. For example, if we take the "20/80 rule," the expected cost with the rule is higher than that with the VoI approach by 10.4 times C_I.

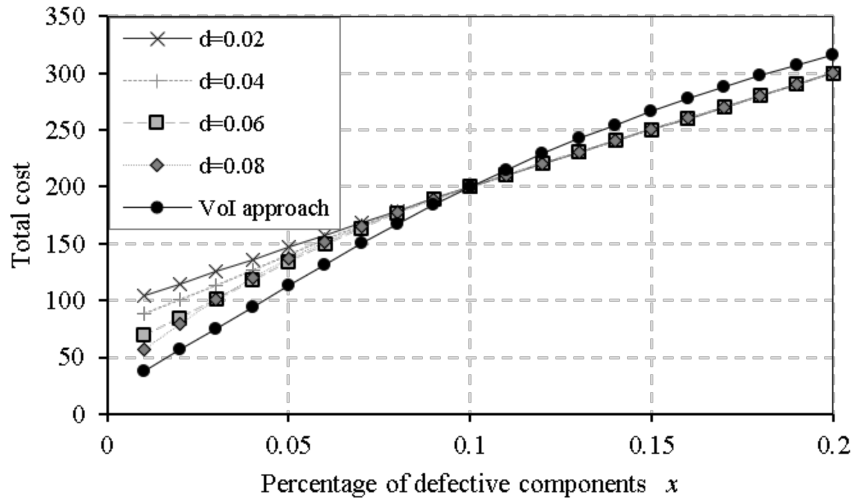

Figure 3.12: Total costs for each approach for $x = 0.01$ to $x = 0.20$.

Table 3.2: Expected total cost for each approach.

Approach		Expected total cost	Sample size
Hypothesis-testing	d=0.02	203.6	93
approach	d=0.04	200.4	76
	d=0.06	197.1	55
	d=0.08	196.6	38
VoI-based approach		193.2	20

Qualitative Comparison

The VoI approach is more suitable for the stated problem than the hypothesis-testing approach, and the limitations of the hypothesis-testing approach are apparent. When we consider fewer failure cases such as $x = 0.01$, we cannot set a reasonable gray region. In addition, when x is too small, approximation using normal distribution for binomial distribution is inappropriate.

3.6 Summary

Diagnostic inspections are carried out to evaluate the condition of a large population of (statistically) similar components found in an engineering system like a power plant or processing unit. There are two possible actions for the engineer. Either inspect every component and replace potentially defective components, or take no action and deal with component failures as they occur. The problem can be formulated as a statistical decision problem in which the information collected by the inspection of a relatively small sample can play a key role.

This chapter presents a systematic VoI approach to determine the optimal sample size as a function of consequential costs associated with the two actions. The comparison between the VoI-based and a traditional approach shows that the VoI approach is preferable to the traditional approach of statistical hypothesis-testing.

The major insights found for diagnostic inspection problems are summarized as follows:

- The more vague and difficult terminal decision-making with only prior information is, the more $ENGS(n^o)$ a decision maker obtains; in that situation, additional information has more value for a decision maker;
- The most sensitive parameter in the diagnostic inspection problem is the break-even value, x_b, which represents the balance of costs: inspection, replacement, and failure costs;
- The highest VoI is expected when the break-even value, x_b, and the mean of prior distribution, $\bar{x}'$, are the same, in which case, the decision maker has the highest risk of taking an inappropriate action based only on prior information;
- The VoI approach is economically more effective unless prior information is irrelevant;
- The VoI approach is more advantageous when the decision maker sets a prior distribution conservatively.

Chapter 4

Inspection Problem with the Random Rate Degradation Model

In this chapter, the value of information (VoI)-based sample size determination (SSD) method is modelled with one of the simplest and the most typical random variable models: the random rate model. This model, which assume a linear degradation process, is first introduced, and then mathematically formulated in the context of a sequential decision-making problem. A numerical example is demonstrated, and insights about using SSD with random variable models are discussed.

4.1 Problem Definition

Consider a population of N statistically identical components in an engineering system, which could be vulnerable to some degradation process and fail if not replaced in time. The system is known to be decommissioned at t_2, and operators intend to inspect each component with cost of C_I and replace them if needed at an earlier time $t_1 < t_2$. A component would break and result in a cost, C_F, if its level of degradation reaches the safety limit, ρ_F. Broken components will be replaced immediately after failure. During an inspection, a component can be preventively replaced at a cost, C_P, such that $C_P < C_F$.

For better estimation of the parameter and prediction of the components' states, $X(t_2)$, inspection will be held at t_1 on selected components. Thus, the problem can be defined as a two-stage decision-making problem. At the first stage, size of sampling inspection is determined, and at the second stage, based on the inspection outcome of $X(t_1)$, operators will decide whether to replace individual components, either the inspected or other, un-inspected components. Note that we assume inspection and replacement actions take no time, and these two stages are in t_1. What operators have to do first is to determine which components will be inspected. For simpler notation, we use X_1 and X_2 instead of $X(t_1)$ and $X(t_2)$, respectively.

The major concern of the operators are the risk of a component breaking between t_1 and t_2, $\mathbb{P}[X_2 > \rho_F \mid X_1 < \rho_F]$. Although the sampling inspection at t_1 costs C_I, it also reduces uncertainty in $\mathbb{P}[X_2 > \rho_F \mid X_1 < \rho_F]$, and consequently, reduces the probabilities of taking an improper action such as replacing safe components.

The assumptions we set in this stated problem are summarized as follows:

- Degradation of components follows a random rate model;
- Components are statistically independent of one another;
- The probability of a replaced component failing before t_2 is negligible;
- Every component survives until t_1 ($x_1 < \rho_F$).

When the nominal life of a component, t_1, is longer than the remaining system life, $t_2 - t_1$, the third assumption is reasonable. The fourth assumption is set to avoid unnecessary calculations in the analysis. The cost arising before t_1 cannot be reduced and is the same for any consequence because operators no option prior to t_1.

4.1.1 Two-Stage Decision-Making Problem

Operators have to decide whether to inspect a component and whether to replace it. At t_1, this problem can be formulated as a sequential decision involves two decisions: inspection and replacement. Figure 4.1 illustrates the two-stage decision-making for a single component inspection problem. At the inspection stage, the operators' options are

$$e(t_1) = \begin{cases} e_1: & Inspect \\ e_2: & Not\ inspect. \end{cases} \tag{4.1}$$

The two actions at the replacement stage for each component are defined as

$$a(t_1) = \begin{cases} a_1: & Replace \\ a_2: & Do\ not\ replace. \end{cases} \tag{4.2}$$

The probability of a component breaking between t_1 and t_2, $\mathbb{P}\left[X_2 > \rho_F \mid X_1 < \rho_F\right]$, will also be evaluated, based on its estimated state for time t_2. Note that X_i represents a component's condition at t_i. Although the sampling inspection at t_1 costs C_I, it also reduces uncertainty in $\mathbb{P}\left[X_2 > \rho_F \mid X_1 < \rho_F\right]$, and consequently, reduces the probabilities of taking an improper action such as replacing safe components.

4.2 Value of Information Analysis

The VoI approach based on random rate model is derived in this section. Because of the assumption for the degradation model, a future condition is precisely predictable once the component is inspected.

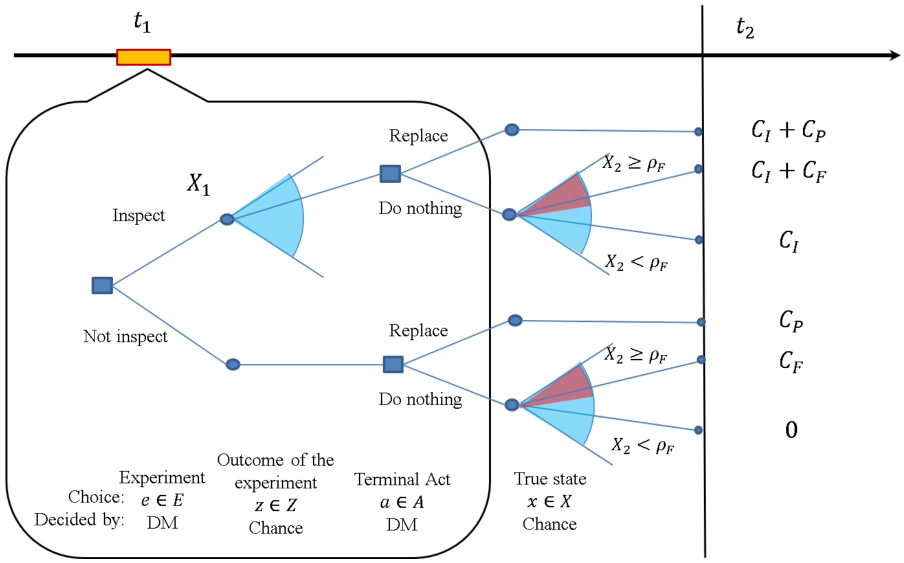

Figure 4.1: Extended decision tree for a single-component inspection problem

4.2.1 Random Rate Model

A random rate model is the simplest case of the random variable model, which is widely used for maintenance decision problems (e.g., Pandey (1998); Stewart and Rosowsky (1998); Hong (2000); Pandey et al. (2009)). Under the model, once the corrosion rate for a component is observed, its future degradation level is precisely predictable. Thus, whether the inspected component will fail before the next diagnostic inspection is obvious. In the random rate model, the deterioration is assumed to linearly proceed over time with a random degradation rate R. The degradation level is formulated as

$$X(t) = Rt. \tag{4.3}$$

The degradation has always non-negative increments; R is assumed to follow a gamma distribution with the probability density function expressed as follows:

$$f_R(r \mid \mu) = ga(r; 1/\nu^2, 1/\mu\nu^2). \tag{4.4}$$

Note that, with the random rate model, $X(t_2)$ of a component can be estimated without any uncertainties once R for the component is revealed through observation of $X(t_1)$.

All components can be categorized to two conditions: $r < r_{F2}$ and $r_{F2} \leq r < r_{F1}$. Note that the condition of $r_{F1} \leq r$ is excluded from the analysis because of the assumption that the components survive at t_1. The two thresholds, r_{F1} or r_{F2} is the degradation rate that reaches ρ_F at t_1 or t_2, respectively. In maintenance decision, r_{F2} becomes a key decision criterion about the inspected outcome r. Note that the criterion satisfies $F_R(r_{F2}) = F_{X_2}(\rho_F)$ and is defined as

$$r_{F2} = \frac{\rho_F}{t_2}. \tag{4.5}$$

4.2.2 Probability Density Functions of Random Variables

The operators' knowledge about μ is summarized as its prior distribution. For simplicity, we take an inverse-gamma distribution as it is a conjugate distribution for a gamma distribution.

$$\begin{aligned} f_M(\mu) &= Iga(\mu; \alpha, \beta) \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\mu}\right)^{\alpha+1} \exp\left(-\frac{\beta}{\mu}\right). \end{aligned} \tag{4.6}$$

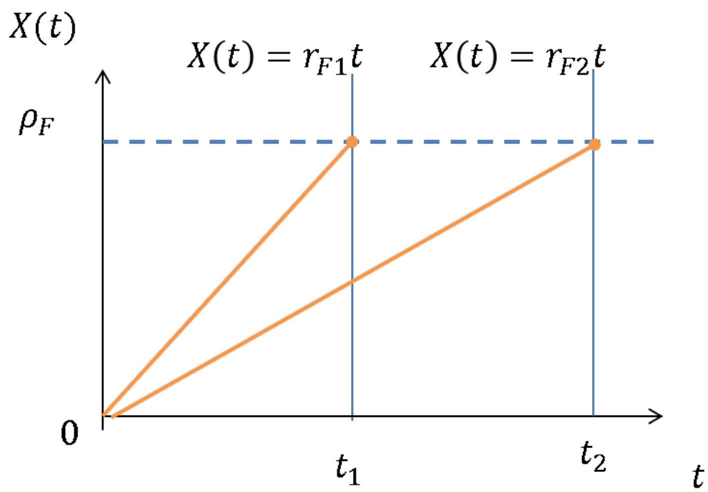

Figure 4.2: Two decision thresholds, r_{F1} and r_{F2}, in the random rate model

where α and β are coefficients selected based on prior information. With inspection outcomes of n components at t_1, the posterior distribution of M becomes

$$f_M(\mu \mid s_n) \quad = \quad Iga\left(\mu; \alpha + \frac{n}{\nu^2}, \beta + \frac{s_n/t_1}{\nu^2}\right). \tag{4.7}$$

Note that the posterior distribution is still inverse-gamma distribution, which is the conjugate distribution for the gamma distribution. With the posterior distribution, we can calculate the expected costs of un-inspected components.

For convenience, Equation (4.4) is rewritten as

$$f_R(r \mid \mu) = ga(r; 1/\nu^2, 1/\mu\nu^2). \tag{4.8}$$

Based on the equation, the probability density function of x_1 is derived as

$$
\begin{aligned}
f_{X_1}(x_1 \mid \mu) &= f_R(r \mid \mu)\left|\frac{dr}{dx_1}\right| \\
&= ga(x_1/t_1; 1/\nu^2, 1/\mu\nu^2) \cdot 1/nt_1 \\
&= ga(x_1; 1/\nu^2, 1/\mu t_1 \nu^2). \qquad (4.9)
\end{aligned}
$$

Because of the additivity nature of a gamma distribution, the probability density function of S_n becomes

$$f_{S_n}(s_n \mid \mu) = ga(s_n; n/\nu^2, 1/\mu t_1 \nu^2). \qquad (4.10)$$

The unconditional distribution of S_n is derived as

$$f_{S_n}(s_n) = \int_0^\infty f_{S_n}(s_n \mid \mu) f_M(\mu) d\mu. \qquad (4.11)$$

The probability density function of r unconditional on μ is derived as

$$f_R(r) = \int_0^\infty f_R(r \mid \mu) f_M(\mu) d\mu. \qquad (4.12)$$

4.2.3 Prior Analysis

Operators, based on the prior distribution, have to assign an expected cost for evaluating and estimating the risk between t_1 and t_2 as follows:

$$\mathbb{E}_M\left[C(a, M)\right] = \begin{cases} C_P & \text{if } a = a_1 \\ C_F \mathbb{E}_M\left[\mathbb{P}\left[R \geq r_{F2} \mid R < r_{F1}\right]\right] & \text{if } a = a_2. \end{cases} \qquad (4.13)$$

Note that $\mathbb{P}[R \geq r_{F2} \mid R < r_{F1}]$ is calculated as

$$\mathbb{P}[R \geq r_{F2} \mid R < r_{F1}] = \frac{F_R(r_{F1}) - F_R(r_{F2})}{F_R(r_{F1})}, \tag{4.14}$$

where $F_R(r)$ denotes the cumulative density function of $f_R(r \mid \mu)$. Comparing the two expected costs for the two actions, operators will take the one that leads to lower cost. The expected optimal cost without inspection is calculated thus:

$$\begin{aligned} C^o_{prior} &= \min_a \mathbb{E}_M [C(a, M)] \\ &= \min (C_P, C_F \mathbb{E}_M [\mathbb{P}[R \geq r_{F2} \mid R < r_{F1}]]) . \end{aligned} \tag{4.15}$$

That selected minimum expected cost is the baseline for evaluating how much a sampling inspection contributes to the effectiveness of an operators' decision.

4.2.4 Inspected Components

Since the random rate model is a deterministic process, whether the component fails depends on whether the observation, r, is greater than r_{F2}. Thus, operators do not have to use the posterior distribution of μ and calculate the expected posterior costs. Conditional on the situation of r, the posterior costs with optimal action become

$$\begin{aligned} C(a^o, M \mid x_1 \geq r_{F2} t_1) &= C_I + \min (C_P, C_F) \\ &= C_I + C_P, \end{aligned} \tag{4.16}$$

$$\begin{aligned} C(a^o, M \mid x_1 < r_{F2} t_1) &= C_I + \min (C_P, 0) \\ &= C_I. \end{aligned} \tag{4.17}$$

Note that operators can always avoid failure of the inspected component because that the inspection outcome add certainty on the operators' decision, hence the analysis for inspected components excludes the cost of failure, C_F, in the formula. Operators will take an action that leads to lower cost than another; the minimum cost is

$$\begin{aligned} C^o_{insp} &= \min_a C(a, M \mid x_1) \\ &= \begin{cases} C_I + C_P & \text{if } r \geq r_{F2} \\ C_I & \text{if } r < r_{F2} \end{cases} . \end{aligned} \tag{4.18}$$

The optimal expected pre-posterior cost for an inspected component is calculated as follows:

$$\begin{aligned} \mathbb{E}_R\left[C^o_{insp}\right] &= C_I + C_P \frac{\int_{r_{F2}}^{r_{F1}} f_R(r)dr}{\int_0^{r_{F1}} f_R(r)dr} \\ &= C_I + C_P \frac{\int_0^{\infty} [F_R(r_{F1}) - F_R(r_{F2})] f_M(\mu)d\mu}{\int_0^{\infty} F_R(r_{F1}) f_M(\mu)d\mu} \end{aligned} \tag{4.19}$$

For the VoI analysis for a single-component problem, the *ENGS* is calculated as the gap between the two expected costs, shown in Equations 4.15 and 4.19, as follows:

$$ENGS_{single} = C^o_{prior} - \mathbb{E}_R\left[C^o_{insp}\right]. \tag{4.20}$$

Note that this *ENGS* represents the net benefit of the perfect information about the future state for a single-component. No influence of the parameter uncertainty in the analysis for a single-component.

4.2.5 Un-inspected Components

For a multiple-component problem, the influence of the parameter uncertainty on un-inspected components need to be identified. Although an un-inspected component has no obtaining information for its own, operators can improve the estimation of its condition at t_2 through reducing parameter uncertainty. With the posterior distribution in Equation (6.20), the expected cost for an un-inspected component is derived as

$$\begin{aligned} C_{non}(a) &= \mathbb{E}_{M|s_n}\left[C(a, M \mid e_2, n, s_n)\right] \\ &= \begin{cases} C_P & \text{if } a = a_1 \\ C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[X_2 > \rho_F \mid X_1 < \rho_F\right]\right] & \text{if } a = a_2, \end{cases} \end{aligned} \tag{4.21}$$

Similar to Equation (3.8), we define the break-even value for the sum of the observation, s_{nb}, as it satisfies the following equation:

$$\mathbb{E}_{M|s_{nb}}\left[\mathbb{P}\left[R \geq r_{F2} \mid R < r_{F1}\right]\right] = \frac{C_P}{C_F}. \tag{4.22}$$

By using the break-even value as a decision criterion, we can obtain the minimum expected cost as follows:

$$\begin{aligned} C^o_{non} &= \min_a C_{non}(a) \\ &= \begin{cases} C_P & \text{if } s_n > s_{nb} \\ C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[R \geq r_{F2} \mid R < r_{F1}\right]\right] & \text{if } s_n \leq s_{nb} \end{cases}. \end{aligned} \tag{4.23}$$

Thus, the expected pre-posterior cost for an un-inspected component is calculated as follows:

$$\begin{aligned} \mathbb{E}_{S_n}\left[C^o_{non}\right] &= C_F \int_0^{s_{nb}} \int_0^{\infty} \mathbb{P}\left[R \geq r_{F2} \mid R < r_{F1}\right] f_M(\mu \mid s_n) f_{S_n}(s_n) d\mu ds_n \\ & \quad + C_P \int_{s_{nb}}^{\infty} \int_0^{\infty} f_M(\mu \mid s_n) f_{S_n}(s_n) d\mu ds_n. \end{aligned} \tag{4.24}$$

4.2.6 Optimal Decision for Inspection

The *ENGS* is defined as the difference between prior and pre-posterior expected costs, represented respectively, as

$$C_{prior}(n) = N \min_a \mathbb{E}_M \left[C(a, M \mid e_2)\right] \tag{4.25}$$

$$C_{prepost}(n) = n\mathbb{E}_R\left[C^o_{insp}\right] + (N-n)\mathbb{E}_{S_n}\left[C^o_{non}\right]. \tag{4.26}$$

We can obtain the $ENGS(n)$ thus:

$$ENGS(n) = C_{prior} - C_{prepost} \tag{4.27}$$

Then, we can derive the optimal sample size, n^o, with which operators will obtain the highest *ENGS*, as follows:

$$n^o = \arg\max_n ENGS(n). \tag{4.28}$$

4.3 Numerical Example

The initial settings are imposed as $N = 100$, $C_F = 100$, $C_P = 10$, $C_I = 1$, $t_1 = 25$, $t_2 = 30$, $1/\nu^2 = 9$, and $\rho_F = 3.0$. The prior distribution for μ is given as $Iga\left(\mu; 540, 47.16\right)$, derived from observed data. The mean and the coefficient of variance (COV) are 0.087 and 0.043, respectively.

4.3.1 Single-Component Problem

The optimal action of the prior analysis is "replace," with its expected cost of 10. The pre-posterior cost is derived as 2.94. The optimal decision is to replace the component if the observed degradation rate, r, is greater than the decision-criterion, $r_{F2} = 0.1$. Then, the *ENGS*, which is the gap between the two costs, is 7.04.

The impacts of the replacement and failure costs are is analysed. Figure 4.3 shows the *ENGS* for each possible combination of C_F and C_P. It indicates that the *ENGS* is proportional to C_P/C_F; as C_F increases or C_P decreases, operators are expected to obtain higher *ENGS*. The *ENGS* has its peak when C_P/C_F is around 20% and drops off as C_P/C_F becomes smaller than that peak value. Since $\mathbb{P}\left[R \geq r_{F2} \mid R < r_{F1}\right]$ with the initial settings is 0.196, when C_P/C_F is around the value, operators cannot make their decision with confidence relying only on their current information.

4.3.2 Multiple-Component Problem

The *ENGS* for each possible sample size is numerically obtained for a multiple-component inspection problem with a random rate model. We use the same initial settings. Without inspection, the optimal action for the prior analysis is a_2, "do-not-replace." The *ENGS* is derived as shown in Figure 4.4. The optimal sample size is 100. For any sample size, most of the *ENGS* is obtained from inspected-components. The percentage of the benefit from un-inspected components is at most 5 % of the *ENGS*. Thus, the optimal sample size is determined based on the balance between inspection and replacement costs, C_I and C_P, respectively.

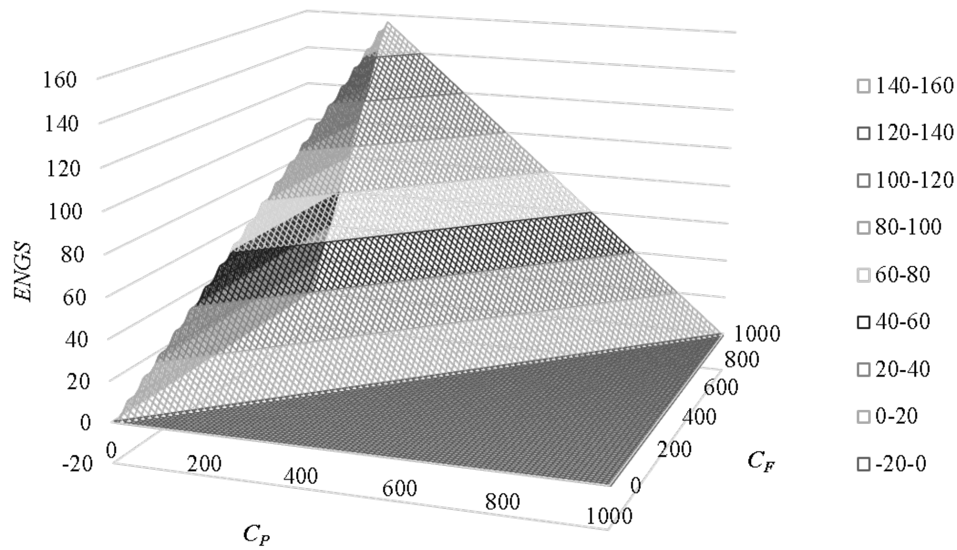

Figure 4.3: *ENGS* for each possible combination of C_F and C_P for a single-component problem with a random rate model ($C_I = 1$)

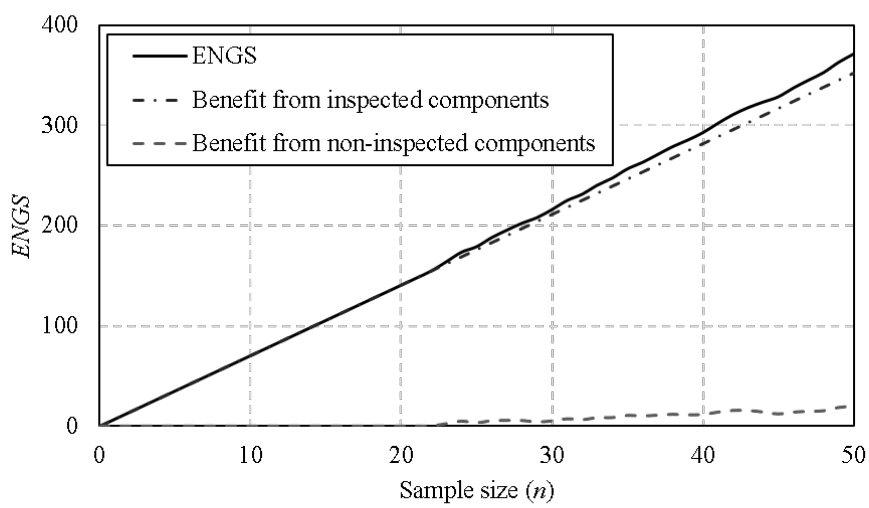

Figure 4.4: *ENGS* and its origin at each sample size ($C_I = 1$, $C_P = 10$, and $C_F = 100$)

4.4 Summary

This chapter has presented a VoI-based SSD method for maintenance problems with the random rate degradation model. The procedure is mathematically formulated in the structure of a two-stage decision-making problem. The formulated mathematical method is applied to a numerical example.

Numerical example illustrated several findings for SSD strategies as follows:

- The benefit obtained by reducing parameter uncertainty is limited;
- No optimal sample size is found (zero or the population size is suggested);
- Because of the nature of the random rate model, an inspection outcome provides perfect information about the future state of the inspected component.

Chapter 5

Inspection Problem with the Gamma Process Degradation Model: Without Parameter Uncertainty

This chapter provides a simple but flexible mathematical sample size determination (SSD) model for a multiple-component system undergoing a realistic degradation process, the gamma process model. The proposed model illustrates how reducing temporal uncertainty contributes to benefiting from inspections even without epistemic uncertainties. The model provides the basis of value of information (VoI)-based SSD approach on condition-based maintenance. Under an assumption of statistical independence among component degradation levels, the model is applied to a practical case study and demonstrates how the *ENGS* is evaluated for both previously inspected and un-inspected components. The evaluation indicates prioritization rule for components for which previous inspection data is available.

5.1 Problem Definition

The same problem introduced in Chapter 4 is adopted except the degradation model. The gamma process model is applied to the VoI-based SSD method instead of the random rate model. In this chapter, no parameter uncertainty is included in the analysis; the components are statistically independent. The assumptions we set in this stated problem are:

- Degradation of components follows the gamma process model;
- Components are statistically independent of one another;
- The probability of a replaced component failing before t_2 is negligible;
- Every component survives until t_1 ($x_1 < \rho_F$).

When the nominal life of a component, t_1, is longer than the remaining system life, $t_2 - t_1$, the third assumption is reasonable.

5.1.1 Two-Stage Decision-Making Problem

The two-stage decision-making for a single-component inspection problem introduced in Chapter 4 is used (see Figure 4.1). The degradation of components is assumed to follow a stochastic degradation model, the gamma process. Operators have prior information about a parameter of the process; however, it is desired to update it after new data become available from inspection. At the inspection-decision stage, the operators take one of the following two inspection options:

$$e(t_1) = \begin{cases} e_1 : & \text{Inspect} \\ e_2 : & \text{Not inspect.} \end{cases} \tag{5.1}$$

Consequently, for each component at t_1, the operators make a decision between the following two actions:

$$a(t_1) = \begin{cases} a_1 : & \text{Replace} \\ a_2 : & \text{Do not replace.} \end{cases} \tag{5.2}$$

5.1.2 Gamma Process Model

Let us set a random variable, $X(t)$, as a degradation level. If the random variable follows a gamma process, it has properties as follows:

- $X(0) = 0$ with probability one;
- $\Delta X(t) = X(t + \Delta t) - X(t) \sim ga(\Delta x; \frac{\Delta t}{\nu^2}, \frac{1}{\mu\nu^2})$ for any $t \geq 0$ and $\Delta t \geq 0$; and
- For any choices of $n \geq 1$ and $0 \leq t_0 < t_1 < \cdots < t_n < \infty$, the random variables $X(t_0), X(t_1) - X(t_0), \cdots, X(t_n) - X(t_{n-1})$ are independent,

where $ga(\Delta x; a, b)$ represents a gamma probability density function with coefficients of a and b; μ and ν are the mean and coefficient of variance of deterioration in a unit time, respectively. The distribution of increments within Δt follows a gamma distribution with the probability density function expressed as

$$g(\Delta x) = \frac{\left(\frac{1}{\mu\nu^2}\right)^{\frac{\Delta t}{\nu^2}}}{\Gamma\left(\frac{\Delta t}{\nu^2}\right)} \Delta x^{\frac{\Delta t}{\nu^2} - 1} \exp\left(-\frac{\Delta x}{\mu\nu^2}\right). \tag{5.3}$$

Then, the cumulative density function of the gamma distribution is defined as

$$G(\rho) = \int_0^{\rho} f_{\Delta X}(\Delta x \mid \mu) d\Delta x. \tag{5.4}$$

5.2 Value of Information Analysis

This section examines the independent component inspection problem, which helps us to understand the nature of the two-stage decision-making process. The expected net gain of sampling (*ENGS*) is calculated and evaluated.

5.2.1 Random Variables

The vector of random variables, $\boldsymbol{Z}_1 = (X_1, X_2)$, is the source of uncertainties. The *ENGS* is obtained by reducing these uncertainties through observing $\boldsymbol{Z}_1$. X_2 is dependent on the deterioration level at t_1 (X_1) as $X_2 = X_1 + \Delta X$. Under the assumption of $x_1 < \rho_F$, the probability density functions of each random variable are as follows:

$$f_{X_1}(x_1) = \frac{g(x_1)}{\int_0^{\rho_F} g(x_1)dx_1}, \tag{5.5}$$

$$f_{X_2|X_1}(x_2 \mid x_1) = g(x_2 - x_1), \tag{5.6}$$

The joint distribution of X_1 and X_2 is

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= f_{X_2|x_1}(x_2 \mid x_1) f_{X_1}(x_1) \\ &= \frac{g(x_1)g(x_2 - x_1)}{\int_0^{\rho_F} g(x_1)dx_1} \end{aligned}$$

$$= \frac{\left(\frac{1}{\mu\nu^2}\right)^{\frac{t_2}{\nu^2}} (x_2 - x_1)^{\frac{\Delta t}{\nu^2}-1} x_1^{\frac{t_1}{\nu^2}-1}}{\Gamma\left(\frac{\Delta t}{\nu^2}\right)\Gamma\left(\frac{t_1}{\nu^2}\right)\int_0^{\rho_F} g(x_1)dx_1} \exp\left(-\frac{x_2}{\mu\nu^2}\right). \tag{5.7}$$

The cumulative density functions of X_1 and X_2 are

$$F_{X_1}(\rho) = \int_0^{\rho} f_{X_1}(x_1)dx_1, \tag{5.8}$$

$$F_{X_2|x_1}(x_2 \mid x_1)(\rho) = \int_{x_1}^{\rho} f_{X_2|x_1}(x_2 \mid x_1)dx_2 = \int_0^{\rho - x_1} g(\Delta x)d\Delta x. \tag{5.9}$$

5.2.2 Structure of Problem

Each possible consequence is evaluated based on its total cost, including any inspection, preventive replacement, and corrective replacement costs. The total cost is function of a replacement action, A, and random variables, $\boldsymbol{Z}_1$. Thus, although operators cannot directly observe X_2, they can have better estimation of X_2 by updating μ knowing perfect information of X_1. The total cost is defined as follows:

$$C(a, \boldsymbol{Z}_1) = \begin{cases} C_P & \text{if } a = a_1 \\ C_F & \text{if } a = a_2 \text{ and } X_2 = x_2 \geq \rho_F \\ 0 & \text{if } a = a_2 \text{ and } X_2 = x_2 < \rho_F. \end{cases} \tag{5.10}$$

To make the problem even simpler, let us consider the expected costs in t_1 and t_2. We get the costs as

$$\begin{aligned} C(a) &= \mathbb{E}_{\boldsymbol{Z}_1}\left[C(a, \boldsymbol{Z}_1)\right] \\ &= \begin{cases} C_P & \text{if } a = a_1 \\ C_F \mathbb{P}\left[X_2 \geq \rho_F\right] & \text{if } a = a_2 \end{cases}. \end{aligned} \tag{5.11}$$

Note that $\mathbb{P}[X_2 \geq \rho_F]$ is calculated as

$$\begin{aligned}\mathbb{P}[X_2 \geq \rho_F] &= \mathbb{P}[X_1 + \Delta X \geq \rho_F] \\ &= \int_0^{\rho_F} [1 - F_{\Delta X}(\rho_F - x_1)] f_{X_1}(x_1) dx_1,\end{aligned} \tag{5.12}$$

where $F_{\Delta X}(\rho_F - x_1)$, which is a cumulative probability density, denotes the probability that the increments is less than $\rho_F - x_1$ so that the component will not fail.

5.2.3 Problem Classification

The stated maintenance problem is classified as decision-making under imperfect information. Although perfect information of X_1 is available, uncertainty still remains on X_2 and consequently on $\boldsymbol{Z}_1$. Thus, since perfect information can be considered as a special case of imperfect information case, we define the *EVSI* and *ENGS* for the problem. However, further discussion of its classification helps to understand the characteristics of the problem. Because of the perfect information assumption on X_1, the problem can be classified as a special case of the expected value of partial perfect information (*EVPPI*). The *EVPPI* is an expected value under perfect information on a subset of random variables. If dependence exists between the random variable in the subset and other random variables that have no direct observations, calculation of *EVPPI* becomes computationally expensive (Claxton and Sculpher, 2006; Jalal et al., 2015). Since X_2 is dependent on X_1, the original assumption breaks and this stated problem can be considered as the *EVPPI* with dependence between random variables with and without perfect information. Chapter 5 defines the problem that is classified as the *EVPPI* with dependent random variables but not requires complicated computational calculation, which will be discussed in Chapter 6.

5.2.4 Prior Analysis

For the prior analysis, we focus on the replacement stage with the given inspection action, "not-inspect (e_2)." Comparing the two expected costs for the two actions in Equation (5.11), operators will take the one that leads to lower cost. The expected optimal cost without inspection is calculated thus:

$$\begin{aligned}\min_a C(a) &= \min(C_P, C_F \mathbb{P}[X_2 > \rho_F]) \\ &= \min(C_P, C_F(1 - F_{X_2}(\rho_F)))\end{aligned} \tag{5.13}$$

That selected minimum expected cost is the baseline for evaluating how much a sampling inspection contributes to the effectiveness of a operators' decision.

5.2.5 Posterior Analysis

For the posterior and pre-posterior analyses, we consider the case with the given inspection action, "inspect (e_1)." At time t_1, the inspection is done and the actual state of the deterioration is revealed. Suppose $X_1 = x_1 < \rho_F$. The expected maintenance cost of an inspected component, which is different from the one without inspection, is

$$\begin{aligned}C(a \mid x_1) &= \mathbb{E}_{\boldsymbol{Z}_1 \mid x_1}[C(a, \boldsymbol{Z}_1 \mid x_1)] \\ &= \begin{cases} C_I + C_P & \text{if } a = a_1 \\ C_I + C_F[1 - F_{\Delta X}(\rho_F - x_1)] & \text{if } a = a_2 \end{cases}.\end{aligned} \tag{5.14}$$

where $[1 - F_{\Delta X}(\rho_F - x_1)]$ is the probability of failure at t_2. The probability is calculated

$$\begin{aligned} 1 - F_{\Delta X}(\rho_F - x_1) &= \mathbb{P}\left[\Delta X > \rho_F - x_1\right] \\ &= \mathbb{P}\left[X_1 + \Delta X \geq \rho_F \mid X_1 = x_1\right]. \end{aligned} \tag{5.15}$$

Since operators will take the action of lower expected cost, the minimum expected cost is

$$\min_a C(a \mid x_1) = C_I + \min_a \{C_P, C_F[1 - F_{\Delta X}(\rho_F - x_1)]\} \tag{5.16}$$

5.2.6 Pre-posterior Analysis

Remember that operators have not actually obtained the inspection outcome at the time of inspection planning, as it is only available after the inspection is undertaken. We thus have to consider all possible outcomes of the inspection, which follows a gamma distribution. With this, the expected optimal cost, with inspection, is expressed as

$$\mathbb{E}_{X_1}\left[\min_a C(a \mid x_1)\right] = C_I + \min_a \{C_P, \mathbb{E}_{X_1}\left[C_F[1 - F_{\Delta X}(\rho_F - x_1)]\right]\}. \tag{5.17}$$

5.2.7 Optimal Decision for Inspection

As defined in Section 2, the *ENGS* is the difference between the two expected optimal costs shown in Equations (5.13) and (5.17). Thus, we can derive the *ENGS* for this single-component problem:

$$ENGS = \min_a C(a) - \mathbb{E}_{X_1}\left[\min_a C(a \mid x_1)\right] \tag{5.18}$$

When the value is positive, operators should inspect a component at t_1 and consequently choose a_1 if $x_1 > x_{1b}$ or a_2 if $x_1 \leq x_{1b}$. When it is negative, they should take an action based only on prior information.

5.3 Problem with Previous Inspection Data

The proposed approach can be extended for a case in which data from previous components is available. Let us assume that, in the past, operators inspected a component at t_0 and obtained its condition, x_0. Let us set Δt_{01} as $t_1 - t_0$, and $X(\Delta t_{01})$ be the degradation increment within Δt_{01}. Once we replace ρ_F, X_1, and x_1 with $\rho_F - x_0$, $x_0 + X(\Delta t_{01})$, and $x_0 + x_{01}$, respectively, in the equations derived in Section 5.2, we can apply the approach to the case with previous inspection data.

5.3.1 Prior Analysis

In this case, Equation (5.13) is modified as follows:

$$\min_a C(a \mid x_0) = \min(C_P, C_F \mathbb{P}[X_2 \geq \rho_F \mid X(\Delta t_{01}) < \rho_F - x_0]) \tag{5.19}$$

where $\mathbb{P}[X_2 \geq \rho_F \mid X(\Delta t_{01}) < \rho_F - x_0]$ is a probability of failure by t_2 given x_0, with an assumption that the component will not fail until t_1. This probability is calculated as

$$\begin{aligned} \mathbb{P}[X_2 \geq \rho_F \mid X(\Delta t_{01}) < \rho_F - x_0] &= \mathbb{P}[X(\Delta t_{01}) + \Delta X \geq \rho_F - x_0] \\ &= \int_0^{\rho_F - x_0} \int_{\rho_F - x_{01} - x_0}^{\infty} f_{\Delta X}(\Delta x) f_{X(\Delta t_{01})}(x_{01}) d\Delta x dx_{01}, \end{aligned} \tag{5.20}$$

where $f_{X(\Delta t_{01})}(x_{01})$ is the probability density function of the degradation progress within Δt_{01}. That selected minimum expected cost is the baseline for evaluating how much a sampling inspection contributes to the effectiveness of an operators' decision.

5.3.2 Pre-posterior Analysis

Based on Equation (5.17), we calculate the expected optimal cost, with inspection, as

$$\begin{aligned}
& \mathbb{E}_{X(\Delta t_{01})|\Delta t_{01}<\rho_F - x_0}\left[\min_a C(a \mid e_1, x_0, x_{01})\right] \\
= \; & C_I + C_P \int_{x_{1b}-x_0}^{\rho_F - x_0} f_{X(\Delta t_{01})}(x_{01}) dx_{01} \\
& + C_F \int_0^{x_{1b}-x_0} \mathbb{P}\left[X_2 \geq \rho_F \mid X(\Delta t_{01}) < \rho_F - x_0\right] f_{X(\Delta t_{01})}(x_{01}) dx_{01}.
\end{aligned} \tag{5.21}$$

5.3.3 Optimal decision for inspection

The object of the decision-making problem is to determine the best inspection action at t_1. By calculating a gap between the two expected optimal costs shown in Equations (5.19) and (5.21), we can derive the *ENGS*, for the problem with previous inspection data:

$$ENGS = \min_a C(a \mid x_0) - \mathbb{E}_{X(\Delta t_{01})|\Delta t_{01}<\rho_F - x_0}\left[\min_a C(a \mid e_1, x_0, x_{01})\right]. \tag{5.22}$$

When the value is positive, operators should inspect a component at t_1 and consequently choose a_1 if $x_1 = x_0 + x_{01} > x_{1b}$ or a_2 if $x_1 = x_0 + x_{01} \leq x_{1b}$. When it is negative, they should take an action based only on prior information.

5.4 Practical Example – Nuclear Piping Systems

Feeder pipes comprise an important part of the primary heat transport system of a CANDU reactor. They connect to a fuel channel and convey coolant from and to pressure tubes

(Figure 1.1). Wall thinning of the outlet feeders, where flow-accelerated corrosion (FAC) is the degradation mechanism, is a major concern of operators as it may cause heavy water leakage when it exceeds the allowable limit. Due to FAC, fatal accidents have occurred in two nuclear plants: the Surry PWR in the US in 1986 and the Mihama 3 PWR in Japan in 2004.

FAC is a process whereby the protective magnetite layer on carbon steel dissolves due to the flowing coolant (water or wet steam) (Dooley and Chexal, 2000). Since the corrosion rate is controlled by the diffusion of iron through the oxide film, the thickness of which reaches a steady-state, the FAC tends to progress at a constant rate (Garland, 2014). The degradation process can be reliably assumed to follow a gamma process model (Yuan, 2007). To plan and conduct well-organized preventive maintenance of the feeder pipes, operators need to implement inspections, which are very expensive because of the high radiation dose.

The objective of the case study is to evaluate the net gain on sampling (*ENGS*) for each component and to identify which components should be inspected in an inspection outage at t_1. We demonstrate the VoI-based approach using this case study. In the case study, the procedure theoretically developed in this chapter is illustrated with real numbers.

5.4.1 Data

At a site, 61 of 380 components were previously inspected: 50 feeder pipes with two measurements, and 11 pipes with three consecutive measurements. The measured minimum thickness of the fourteen-probe ultrasonic testing bracelet has been stored for each pipe at each inspection outage time. The measurement locations are approximately the same between the outages. The initial wall thickness is estimated at 5.3 mm. The data on the inspected degradation level is illustrated in Figure 5.1. Based on the data, the parameters of the gamma process, μ and ν, are estimated through the maximum likelihood estimation (MLE); the parameters are estimated at $(\mu, \nu) =$

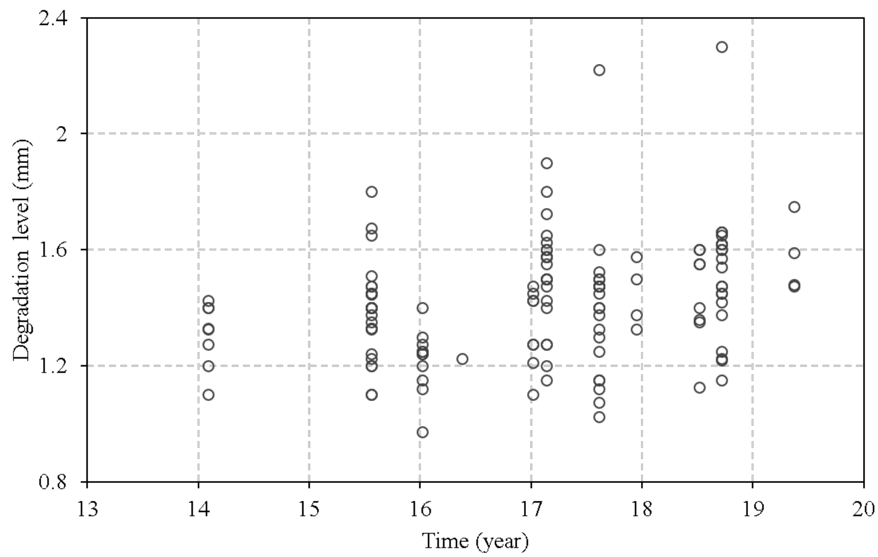

Figure 5.1: Inspected data of 50 feeder pipes with two measurements and 11 pipes with three consecutive measurements

$(0.0836, 0.584)$. We impose several initial settings: $C_F = 100$, $C_P = 10$, $C_I = 1$, $t_1 = 25$, $t_2 = 30$, and $\rho_F = 3.0$.

5.4.2 Single-Component without Previous Inspection

As a first step of a demonstration, we show the VoI-based analysis for a single-component inspection problem step by step. Since the probability of failure between t_1 and t_2 is $\mathbb{P}\left[X_2 \geq \rho_F\right] = 0.0391$, operators can obtain the prior expected cost introduced in Equation (5.13), as

$$C(a) = \begin{cases} 10.0 & \text{if } a = a_1 \\ 3.91 & \text{if } a = a_2 \end{cases}$$

Comparing the two expected costs for the two actions, without inspection, operators will take a_2, "do not replace," with its expected cost 3.91.

For better understanding of pre-posterior analysis, let us consider a fictitious situation where we obtain $X_1 = 2.50$. Based on Equation (5.14), the cost function of an inspected component is

$$C(a \mid x_1) = \begin{cases} 11.0 & \text{if } a = a_1 \\ 23.1 & \text{if } a = a_2 \end{cases}$$

The minimum expected cost is $\min_a C(a \mid x_1) = 11.0$, and the optimal action given $X_1 = 2.50$ is a_1, "replacement." Note that the break-even value of x_1, x_{1b}, is 2.38.

Remember that operators have not actually obtained any sampling inspection outcome. We thus have to consider all possible outcomes and their probability of occurrence. The expected optimal cost with inspection, which is formulated in Equation (5.17), is calculated as

$$\mathbb{E}_{X_1}\left[\min_a C(a \mid x_1)\right] = 2.14,$$

The object of the decision-making problem is to determine the best (most cost effective) inspection at t_1. By calculating a gap between the two expected optimal costs formulated in Equation (6.17), we can derive *ENGS*:

$$\begin{aligned} ENGS &= 3.91 - 2.35 \\ &= 1.56. \end{aligned}$$

Since the value is positive, operators should inspect a component at t_1 and consequently choose a_1 if $x_1 > 2.38$ or a_2 if $x_1 \leq 2.38$.

Table 5.1: Number of components for each inspection decision

Inspection decision	Do not inspect	Inspect
Previously inspected components	46	15
Components without inspection	0	319

5.4.3 Value of Inspection with Previous Observation Data

For each component, we derive $ENGS$, decide if the component should be inspected or reinspected, and summarize the results in Table 5.1. This analysis shows that 334 components have a positive $ENGS$, meaning they should be inspected. Note that 15 are previously inspected components, and all the previously un-inspected components have positive ENGS. Only five components have higher $ENGS$ values than the components that have never been inspected before. This result indicates that inspecting components inspected previously tends to be less of a priority than is examining un-inspected components.

The optimal plan is to inspect 334 out of 380 components; the accumulated $ENGS$ is 512, which is the total contribution of the inspection at t_1. Figure 5.2 illustrates the relation among x_0, t_0, and $ENGS$ for each component. The components for which operators have a low risk to make an incorrect decision have a negative ENGS, and the components where the risk is high have positive ENGS.

5.5 Summary

This chapter has presented a VoI-based SSD method for maintenance problems, one that considers temporal uncertainty in the model. First, we introduced the method with the baseline mathematical model and showed the general characteristics of the stated problem. The model was extended to the case of previous inspection data. We have demonstrated the proposed method

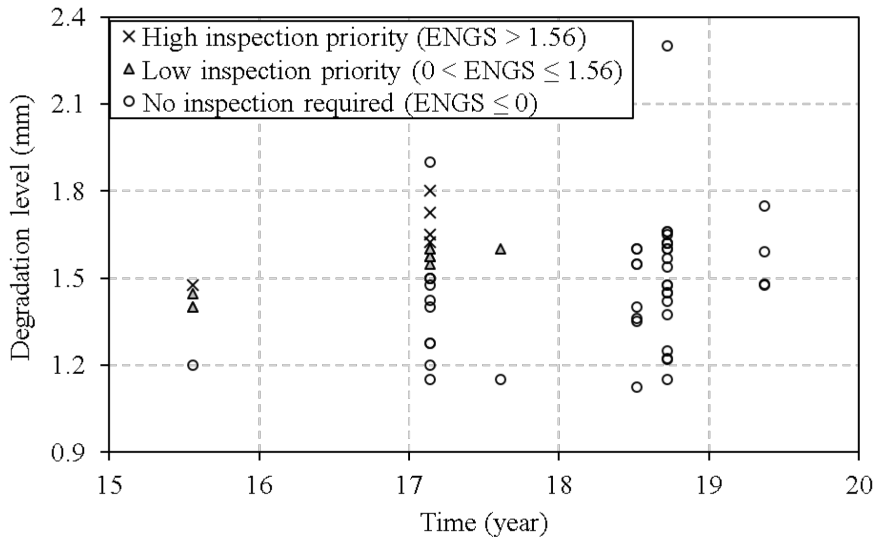

Figure 5.2: *ENGS* for each component with previous inspection data

through a realistic numerical example. The results provide operators with not only the sample size but also the priority of inspection, which is useful if a large sample size is not feasible because of other restrictions such as resources.

Chapter 6

Inspection Problem with Gamma Process Degradation Model: With Parameter Uncertainty

This chapter extends the sample size determination (SSD) model proposed in Chapter 5 for a maintenance and inspection for multiple-components that are dependent each other through the shared parameter uncertainty. Because of the dependency, operators need to take all the conditions of all inspected components in their consideration so that the SSD model used in Chapter 5 cannot be applied to the multiple-component system. The additivity characteristics of the gamma distribution mean that only the sum of the observation outcomes needs to be considered – as a representative variable – instead of considering the conditions of each component separately. This simplification is critical for the updating process in pre-posterior analysis.

Moreover, the impacts of epistemic (parameter) and aleatory (temporal) uncertainties are explicitly compared. These uncertainties have not been comprehensively analysed; the traditional value of information (VoI) concept considers only parameter uncertainties, and the studies with partially observable Markov decision process (POMDP) mainly focus on measurement errors. The

VoI-based approach with a time-dependent degradation process is modelled and quantitatively analysed with both uncertainties included.

6.1 Problem Definition

The same problem used in Chapter 5, which is originally introduced in Chapter 4, is adopted except that the degradation model includes parameter uncertainty in it. Because of the parameter uncertainty, the degradations of components are independent only if the parameters of the degradation model are given. Thus, the degradations are conditional independent with the parameter uncertainty. The assumptions we set in this stated problem are summarized as follows:

- Degradation of components follows the gamma process model;
- Components are statistically independent of one another given parameters of the degradation model;
- The probability of a replaced component failing before t_2 is negligible;
- Every component survives until t_1 ($x_1 < \rho_F$).

When the nominal life of a component, t_1, is longer than the remaining system life, $t_2 - t_1$, the third assumption is reasonable.

6.2 Single-Component Problem

This chapter builds the model with both temporal (aleatory) and parameter (epistemic) uncertainties. Based on the problem with only temporal uncertainty defined in Chapter 5, the single-component problem is first stated for illustrative reason, and the full model is next introduced.

6.2.1 Prior and Posterior Distribution

The prior distribution of M is assumed to follow an inverse-gamma distribution as it is a conjugate distribution for a gamma distribution.

$$\begin{aligned} f_M(\mu) &= Iga(\mu; \alpha, \beta) \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\mu}\right)^{\alpha+1} \exp\left(-\frac{\beta}{\mu}\right). \end{aligned} \tag{6.1}$$

where α and β are coefficients selected based on prior information, which can be previous data or experts' knowledge.

Once inspection outcome is obtained as $X_1 = x_1$, through a Bayesian updating procedure, a posterior distribution of M is obtained:

$$f_M(\mu \mid x_1) = Iga\left(\mu; \alpha + \frac{t_1}{\nu^2}, \beta + \frac{x_1}{\nu^2}\right). \tag{6.2}$$

6.2.2 Random Variables

The vector of the random variables, $\boldsymbol{Z}_2 = (M, X_1, X_2)$, is the source of uncertainties. The *ENGS* is obtained by reducing these uncertainties through observing $\boldsymbol{Z}_2$. Under the assumption of $x_1 < \rho_F$, the conditions at t_1 and t_2 are estimated based on the probability density functions (PDFs) of X_1 and X_2, respectively, as follows:

$$f_{X_1|M}(x_1 \mid \mu) = \frac{g(x_1)}{\int_0^{\rho_F} g(x_1)dx_1}, \tag{6.3}$$

$$f_{X_2|M,X_1}(x_2 \mid \mu, x_1) = g(x_2 - x_1), \tag{6.4}$$

The increment of the degradation level between t_1 and t_2, Δx, has the following PDF:

$$f_{\Delta X|M,X_1}(\Delta x \mid \mu, x_1) \quad = \quad g(\Delta x) \tag{6.5}$$

The joint distribution of M and X_1 is

$$\begin{aligned} f_{M,X_1}(\mu, x_1) &= f_{X_1}(x_1 \mid \mu) f_M(\mu) \\ &= \frac{g(x_1)}{\int_0^{\rho_F} g(x_1) dx_1} Iga(\mu; \alpha, \beta) \\ &= \frac{\left(\frac{x_1}{\nu^2}\right)^{\frac{t_1}{\nu^2}} \beta^\alpha}{x_1 \Gamma(\frac{t_1}{\nu^2}) \Gamma(\alpha) \int_0^{\rho_F} g(x_1) dx_1} \left(\frac{1}{\mu}\right)^{\alpha + \frac{t_1}{\nu^2} + 1} \exp\left(-\frac{1}{\mu}\left(\beta + \frac{x_1}{\nu^2}\right)\right), \end{aligned} \tag{6.6}$$

The marginal distribution of X_1 is derived as

$$\begin{aligned} f_{X_1}(x_1) &= \int_0^\infty f_{M,X_1}(\mu, x_1) d\mu \\ &= \frac{\Gamma(\alpha + \frac{t_1}{\nu^2})}{\Gamma(\frac{t_1}{\nu^2}) \Gamma(\alpha)} \frac{\left(\frac{x_1}{\nu^2}\right)^{\frac{t_1}{\nu^2}} \beta^\alpha}{x_1 \left(\beta + \frac{x_1}{\nu^2}\right)^{\alpha + \frac{t_1}{\nu^2}} \int_0^{\rho_F} g(x_1) dx_1}. \end{aligned} \tag{6.7}$$

Note that the cumulative density function of X_1 is difficult to calculate the general form. To evaluate the integrations, the sample from the marginal distribution is created from the joint distribution, $f_{M,X_1}(\mu, x_1)$, using Gibbs sampling. See Section 6.4.1 for more detail. Another approximation method for calculating the integrations is introduced in Section 6.4.3.

For the posterior analysis, the joint distribution of M and X_2 given $X_1 = x_1$ is

$$
\begin{aligned}
& f_{M,X_2|X_1}(\mu, x_2 \mid x_1) \\
= \; & f_{X_2|M,X_1}(x_2 \mid \mu, x_1) f_M(\mu \mid x_1) \\
= \; & \frac{\left(\frac{x_2-x_1}{\nu^2}\right)^{\frac{\Delta t}{\nu^2}} (\beta + \frac{x_1}{\nu^2})^{\alpha+\frac{t_1}{\nu^2}}}{(x_2-x_1)\Gamma(\frac{\Delta t}{\nu^2})\Gamma(\alpha+\frac{t_1}{\nu^2})} \left(\frac{1}{\mu}\right)^{\alpha+\frac{t_1+\Delta t}{\nu^2}+1} \exp\left(-\frac{1}{\mu}\left(\beta + \frac{x_1+(x_2-x_1)}{\nu^2}\right)\right), \quad (6.8)
\end{aligned}
$$

The marginal distribution of X_2 given $X_1 = x_1$ is

$$
\begin{aligned}
f_{X_2|X_1}(x_2 \mid x_1) &= \int_0^\infty f_{M,X_2|X_1}(\mu, x_2 \mid x_1) d\mu \\
&= \frac{\left(\frac{x_2-x_1}{\nu^2}\right)^{\frac{\Delta t}{\nu^2}} (\beta + \frac{x_1}{\nu^2})^{\alpha+\frac{t_1}{\nu^2}} \Gamma(\alpha + \frac{t_1+\Delta t}{\nu^2})}{(x_2-x_1)(\beta + \frac{x_1+(x_2-x_1)}{\nu^2})^{\alpha+\frac{t_1+\Delta t}{\nu^2}} \Gamma(\frac{\Delta t}{\nu^2})\Gamma(\alpha+\frac{t_1}{\nu^2})}, \quad (6.9)
\end{aligned}
$$

The cumulative density functions (CDFs) of X_1 and X_2 are

$$
F_{X_1|M}(\rho) = \int_0^\rho f_{X_1|M,X_2}(x_1 \mid \mu, x_2) dx_1, \quad (6.10)
$$

$$
\begin{aligned}
F_{X_2|M,X_1}(\rho) &= \int_{x_1}^\rho f_{X_2|M,X_1}(x_2 \mid \mu, x_1) dx_2 = \int_0^{\rho-x_1} g(\Delta x) d\Delta x, \quad (6.11) \\
F_{X_2|X_1}(\rho) &= \int_{x_1}^\rho f_{X_2|X_1}(x_2 \mid x_1) dx_2 \\
&= \int_0^{\rho-x_1} \frac{\left(\frac{\Delta x}{\nu^2}\right)^{\frac{\Delta t}{\nu^2}} (\beta + \frac{x_1}{\nu^2})^{\alpha+\frac{t_1}{\nu^2}} \Gamma(\alpha + \frac{t_1+\Delta t}{\nu^2})}{\Delta x(\beta + \frac{x_1+\Delta x}{\nu^2})^{\alpha+\frac{t_1+\Delta t}{\nu^2}} \Gamma(\frac{\Delta t}{\nu^2})\Gamma(\alpha+\frac{t_1}{\nu^2})} d\Delta x. \quad (6.12)
\end{aligned}
$$

Note that $F_{X_2|X_1}(\rho)$ is the same as $F_{\Delta X}(\rho_F - x_1)$, which is the CDF of Δx.

6.2.3 Prior Analysis

First, as a prior analysis, operators need to know the best possible decisions when no sampling-inspection data is available. The total cost is dependent on an action, A, and random variables, $\boldsymbol{Z}_2 = (M, X_1, X_2)$. Operators, based only on the prior distribution of M, evaluate each action by estimating the risk between t_1 and t_2. Equation (5.11) is modified as follows:

$$\mathbb{E}_{\boldsymbol{Z}_2}\left[C(a, \boldsymbol{Z}_2)\right] = \begin{cases} C_P & \text{if } a = a_1 \\ C_F \mathbb{E}_M\left[\mathbb{P}\left[X_2 > \rho_F\right]\right] & \text{if } a = a_2, \end{cases} \tag{6.13}$$

Note that, for convenience, $x(t_i)$ for $i = 1, 2$ is shortened to x_i. According to the assumption, although operators do not know what X_1 is, we assume that all possible values as a x_1 satisfies $0 < x_1 < \rho_F$.

Comparing the two expected costs for the two actions, operators will take the one that leads to lower cost. The expected optimal cost without inspection is calculated thus:

$$\min_a \mathbb{E}_{\boldsymbol{Z}_2}\left[C(a, \boldsymbol{Z}_2)\right] = \min\left(C_P, C_F \mathbb{E}_M\left[\mathbb{P}\left[X_2 > \rho_F \mid X_1 < \rho_F\right]\right]\right). \tag{6.14}$$

6.2.4 Posterior Analysis

Suppose that operators have obtained an outcome for a sampling inspection, $X_1 = x_1 < \rho_F$. The operators will then choose the best action according to a comparison between the expected costs with updated information about X_2 and M. With the posterior distribution in Equation (6.2), operators will take an action that leads to lower cost than another; the minimum expected cost is

$$\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_1}\left[C(a, \boldsymbol{Z}_2 \mid x_1)\right] \quad = \quad C_I + \min_a \left\{C_P, C_F \mathbb{E}_{M|x_1}\left[1 - F_{\Delta X}(\rho_F - x_1)\right]\right\}. \qquad (6.15)$$

6.2.5 Pre-posterior Analysis

Remember that the inspection outcome, X_1, is a random variable. We calculate the expected optimal cost, with inspection, as

$$\mathbb{E}_{X_1}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_1}\left[C(a, \boldsymbol{Z}_2 \mid x_1)\right]\right] \quad = \quad \frac{\int_0^{\rho_F} \min_a \mathbb{E}_{\boldsymbol{Z}_2|x_1}\left[C(a, \boldsymbol{Z}_2 \mid x_1)\right] f_{X_1}(x_1) dx_1}{\int_0^{\rho_F} f_{X_1}(x_1) dx_1}. \qquad (6.16)$$

By calculating a gap between the two expected optimal costs shown in Equations (6.14) and (6.16), we can derive the *ENGS* for this single-component problem:

$$ENGS \quad = \quad \min_a \mathbb{E}_{\boldsymbol{Z}_2}\left[C(a, \boldsymbol{Z}_2)\right] - \mathbb{E}_{X_1}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_1}\left[C(a, \boldsymbol{Z}_2 \mid x_1)\right]\right] \qquad (6.17)$$

For single-component problem, the decision of inspection is contingent only on the timing of inspection (t_1), if other parameters are kept constant. Alternatively, if t_1 is fixed, then the decision of inspection is dependent on the relative inspection cost.

6.3 Multiple-Component Problem

This section expands the model proposed in Section 6.2 to an N-component system problem by having accumulated the features of the single-component problem. Through the analysis with the model, we can answer what sample size is the best for the group of homogeneous components.

Table 6.1: Prior and pre-posterior analysis for multiple-component problem

		E	Z	A	X
Prior analysis		$n = 0$	No outcomes	$a^{(prior)}$ for all the components	$x_2^{(1)}, x_2^{(2)}, \cdots, x_2^{(N)}$
Pre -posterior analysis	Each inspected component	n	$x_1^{(i)}$ and $s_{n-1}^{(-i)}$	$a^{(i)}$ for each i $(i = 1, 2, \cdots, n)$	$x_2^{(1)}, \cdots, x_2^{(n)}$
	Un-inspected components		s_n	$a^{(non)}$ for all the non-inspected components	$x_2^{(n+1)}, \cdots, x_2^{(N)}$

In the N-components problem, operators can select an inspection sample size, n, which represents the action at the inspection stage, $e(t_1) = n$. Note $n = 0$ means no inspection. By observing the inspection outcomes of n components, $z = (x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)})$, operators can choose a set of actions for the N components, $a(t_1) = (a^{(1)}, a^{(2)}, \cdots, a^{(N)})$. They then determine a consequence for each component based on all the components' states at time t_2, $x = (x_2^{(1)}, x_2^{(2)}, \cdots, x_2^{(N)})$, obtained by following a mathematically modelled stochastic degradation process, which has an unknown parameter, M. As the size of sampling inspection, n, increases, operators can reduce uncertainty in their estimation for X_2 through the observed $x_1^{(i)}$ and updated knowledge about parameter M from $z = (x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)})$. For the n components to be inspected, we consider the upper path of the decision tree in Figure 4.1, and for un-inspected components, we calculate the possible consequences based on the lower path of the decision tree. The problem is summarized in Table 6.1. Note that $a^{(prior)}$, $a^{(i)}$, and $a^{(non)}$ are one of the two actions: replace (a_1) and do-not-replace (a_2).

6.3.1 Random Variables

The random variables that affect the total expected cost are summarized as a vector, $\boldsymbol{Z}_3 = (M, X_1, S_{n-1}, X_2)$. X_1, X_2, and S_{n-1} are conditional independent variables, which are independent if μ is known. Because of the additivity characteristics of the gamma process, the probability density function of S_{n-1} is derived as follows:

$$f_{S_{n-1}}(s_{n-1} \mid \mu) = ga(s_{n-1}; nt_1/\nu^2, 1/\mu\nu^2). \tag{6.18}$$

The probability density function of X_1 and S_{n-1} unconditional on μ is derived as

$$\begin{aligned} f_{X_1,S_{n-1}}(x_1, s_{n-1}) &= \int_0^\infty f_{X_1}(x_1 \mid \mu) f_{S_{n-1}}(s_{n-1} \mid \mu) f_M(\mu) d\mu \\ &= \frac{\Gamma(\alpha + \frac{nt_1}{\nu^2})}{\Gamma(\frac{t_1}{\nu^2})\Gamma(\frac{(n-1)t_1}{\nu^2})\Gamma(\alpha)} \frac{\left(\frac{x_1}{\nu^2}\right)^{\frac{t_1}{\nu^2}} \frac{s_{n-1}}{\nu^2}^{\frac{(n-1)t_1}{\nu^2}} \beta^\alpha}{x_1 s_{n-1} \left(\beta + \frac{x_1+s_{n-1}}{\nu^2}\right)^{\alpha+\frac{nt_1}{\nu^2}}} \end{aligned} \tag{6.19}$$

Figure 6.1 illustrates the joint distribution of X_1 and S_{n-1} when $n = 10$.

With inspection outcomes of n components at t_1, the posterior distribution of M becomes

$$f_M(\mu \mid x_1, s_{n-1}) = Iga\left(\mu; \alpha + \frac{nt_1}{\nu^2}, \beta + \frac{x_1 + s_{n-1}}{\nu^2}\right). \tag{6.20}$$

Note that the posterior distribution is still inverse-gamma distribution, which is the conjugate distribution for the gamma distribution. With the posterior distribution, we can calculate the expected costs of both inspected and un-inspected components.

Based on the posterior distribution, the distribution of ΔX is derived as follows:

$$\begin{aligned} f_{\Delta X|X_1,S_{n-1}}(\Delta x \mid x_1, s_{n-1}) &= \int_0^\infty f_{\Delta X|M}(\Delta x \mid \mu) f_{M|X_1,S_{n-1}}(\mu \mid x_1, s_{n-1}) d\mu \\ &= \frac{\Gamma(\alpha + \frac{nt_1}{\nu^2} + \frac{t_2-t_1}{\nu^2})}{\Gamma(\alpha + \frac{nt_1}{\nu^2})\Gamma(\frac{t_2-t_1}{\nu^2})} \frac{\left(\frac{\Delta x}{\nu^2}\right)^{\frac{t_2-t_1}{\nu^2}} \left(\beta + \frac{x_1+s_{n-1}}{\nu^2}\right)^{\alpha+\frac{nt_1}{\nu^2}}}{\Delta x \left(\beta + \frac{x_1+s_{n-1}+\Delta x}{\nu^2}\right)^{\alpha+\frac{nt_1}{\nu^2}+\frac{t_2-t_1}{\nu^2}}}. \end{aligned} \tag{6.21}$$

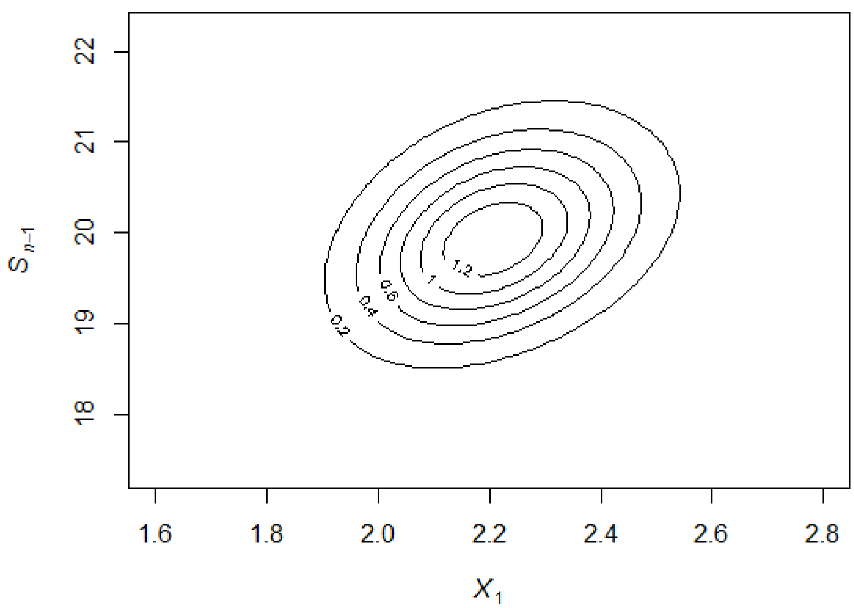

Figure 6.1: Joint distribution of X_1 and S_9 ($n = 10$)

6.3.2 Prior Analysis

The total cost is functional on a replacement action, A, and random variables, $\boldsymbol{Z}_3 = (M, X_1, S_n, X_2)$. In the prior analysis, since $n = 0$, the vector of $\boldsymbol{Z}_3$ becomes the same as $\boldsymbol{Z}_2$, which is used for single-component problem. Since there is no observation results for prior analysis, the optimal decisions for every components are the same. The optimal decisions can be derived by Equation (6.14).

6.3.3 Posterior Analysis

The changes of the model appear when we are updating information about M based on inspection outcomes, $x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)}$. In a general problem, we have to evaluate all possible combinations of each component's state, and doing so is one of the difficulties of applying the VoI concept to maintenance problems. However, with the gamma process model and its conjugate prior distribution, inverse-gamma distribution, we can use a sum of the outcomes, $s_n = x_1^{(1)} + x_1^{(2)} + \cdots + x_1^{(n)} = x_1^{(i)} + s_{n-1}^{(-i)}$, as a representative random variable of all possible outcomes.

With these modified posterior distribution and random variables, we can calculate the pre-posterior costs for both inspected and un-inspected components. By summing up the value for each component, we can derive the *ENGS* of a system.

6.3.4 Pre-posterior Analysis

Since the prior distribution of M is the same as in the single-component problem, and the prior cost for both inspected and un-inspected components is the same as the one defined in Section 6.2.3, for convenience, we rewrite the prior cost as follows:

$$\min_a \mathbb{E}_{\boldsymbol{Z}_3} \left[C(a, \boldsymbol{Z}_3) \right] = \min_a \left\{ C_P, C_F \mathbb{E}_M \left[\mathbb{P} \left[X_2 > \rho_F \right] \right] \right\}. \tag{6.22}$$

For the i^{th} inspected component, the operator makes the replacement decision based on two pieces of evidence. The first one is the actual condition of the component, $x_1^{(i)}$. Without measurement error, this inspection outcome eliminates completely the temporal uncertainty. The second piece, which is indirect, is the sum of conditions of the other $n-1$ components, denoted by $s_{n-1}^{(-i)}$. The sum reduces the parameter uncertainty by contributing to the updating of the model parameter of the gamma process.

Inspected Components

The expected pre-posterior cost for an inspected component is calculated as follows:

$$\begin{aligned} C_{insp}(a) &= \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_1^{(i)}, s_{n-1}^{(-i)})\right] \\ &= \begin{cases} C_I + C_P & \text{if } a = a_1 \\ C_I + C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[\Delta X > \rho_F - x_1^{(i)}\right]\right] & \text{if } a = a_2, \end{cases} \end{aligned} \tag{6.23}$$

After comparing the costs of each action, operators can choose an action that leads to lower expected cost. For a multiple-component problem, we need to consider the two random variables, $X^{(i)}(t_1)$ and $S_{n-1}^{(-i)}$, simultaneously, thus the break-even values, which are criteria for choosing a better action, form a line in a two-dimensional space of $(x_1^{(i)}, s_{n-1}^{(-i)})$, as shown in Figure 6.2. Note that $h_1(x_1^{(i)}, s_{n-1}^{(-i)}) = 0$ represents the break-even line on which the expected costs from the two actions become the same and is formulated as

$$h_1(x_1^{(i)}, s_{n-1}^{(-i)}) = C_P - C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[\Delta X > \rho_F - x_1^{(i)}\right]\right]. \tag{6.24}$$

Then, with the break-even values, we can describe the minimum expected cost as

$$\begin{aligned} C_{insp}^{o} &= \min_a C_{insp}(a) \\ &= C_I + \min_a \left\{C_P, C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[\Delta X > \rho_F - x_1^{(i)}\right]\right]\right\} \\ &= \begin{cases} C_I + C_P & \text{if } h(x_1^{(i)}, s_{n-1}^{(-i)}) < 0 \\ C_I + C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[\Delta X > \rho_F - x_1^{(i)}\right]\right] & \text{if } h(x_1^{(i)}, s_{n-1}^{(-i)}) \geq 0, \end{cases} \end{aligned} \tag{6.25}$$

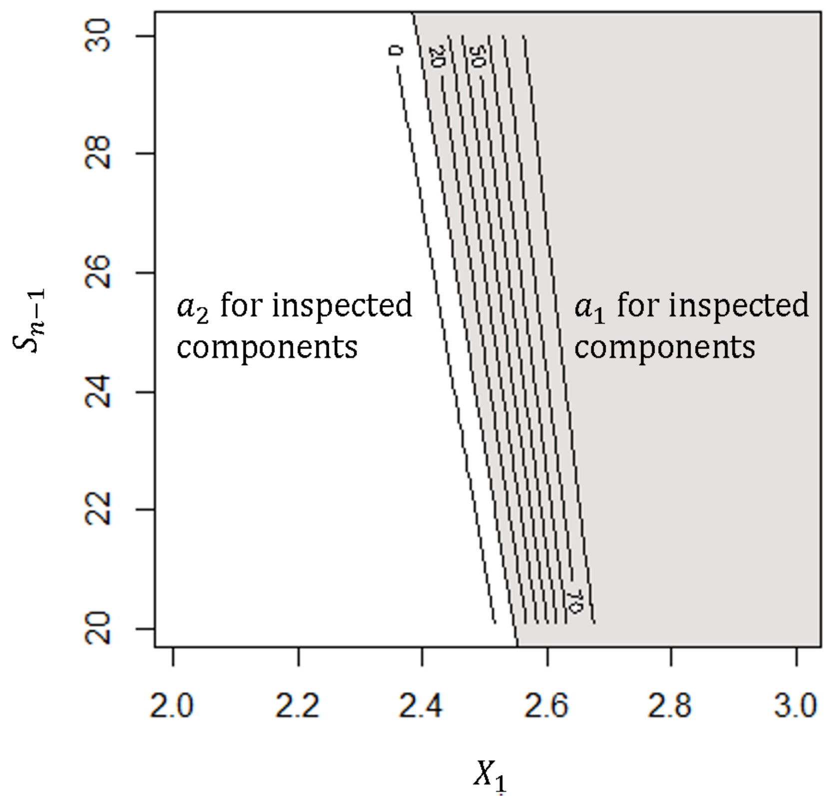

Figure 6.2: Cost of taking a_2 and optimal actions after observing $s_{n-1}^{(-i)}$ and $x_1^{(i)}$ $(n = 10, C_I = 3, C_P = 10, and C_F = 100)$

Un-inspected Components

Although an un-inspected component has no obtaining information for its own, operators can improve the estimation of its condition at t_2 through reducing parameter uncertainty. With the posterior distribution in Equation (6.20), the expected cost for an un-inspected component is derived as

$$\begin{aligned} C_{un}(a) &= \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_n)\right] \\ &= \begin{cases} C_P & \text{if } a = a_1 \\ C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[X_2 > \rho_F\right]\right] & \text{if } a = a_2, \end{cases} \end{aligned} \tag{6.26}$$

Similar to Equation (3.8), we define the break-even value for the sum of the observation, s_{nb}, as it satisfies the following equation:

$$\mathbb{E}_{M|s_{nb}}\left[\mathbb{P}\left[X_2 > \rho_F\right]\right] = \frac{C_P}{C_F}. \tag{6.27}$$

By using the break-even value as a decision criterion, we can obtain the minimum expected cost as follows:

$$\begin{aligned} C_{un}^o &= \min_a C_{un}(a) \\ &= \min_a \left\{C_P, C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[X_2 > \rho_F\right]\right]\right\} \\ &= \begin{cases} C_P & \text{if } s_n > s_{nb} \\ C_F \mathbb{E}_{M|s_n}\left[\mathbb{P}\left[X_2 > \rho_F\right]\right] & \text{if } s_n \leq s_{nb} \end{cases}. \end{aligned} \tag{6.28}$$

6.3.5 Optimal Decision for Inspection

The *ENGS* is defined as the difference between prior and pre-posterior expected costs, represented respectively, as

$$C_{prior}(n) = N \min_{a} \mathbb{E}_{\boldsymbol{Z}_3} \left[C(a, \boldsymbol{Z}_3 \mid e_2) \right] \quad (6.29)$$

$$C_{prepost}(n) = n\mathbb{E}_{X_1^{(i)}, S_{n-1}^{(-i)}} \left[C_{insp}^{o} \right] + (N-n)\mathbb{E}_{S_n} \left[C_{un}^{o} \right] \quad (6.30)$$

We can obtain the *ENGS*(n) thus:

$$ENGS(n) = C_{prior} - C_{prepost} \quad (6.31)$$

Then, we can derive the optimal sample size, n^o, with which operators will obtain the highest ENGS, as follows:

$$n^o = \arg\max_{n} ENGS(n). \quad (6.32)$$

6.4 Computational Algorithms

The calculation of *ENGS* involves the evaluation of several multi-dimensional integrations and minimization operators. For example, in Equation (6.30), the first term involves an integration with respect to $x_1^{(i)}$ and $s_{n-1}^{(-i)}$. The two variables are dependent on each other, unconditional on μ; that is, they are conditional independent. Although the joint probability density function can be analytically derived, large values in gamma functions make numerical calculation difficult. To calculate these multiple-integrations, we employ a numerical calculation method, Markov Chain Monte Carlo (MCMC). How the MCMC simulations are organized with minimization operators is explained in this section. The original algorithms are introduced first, and two approximation methods, memoization and PMF, are adopted to reduce the computational cost.

6.4.1 Gibbs Sampling

With MCMC, a set of random variables can be simulated from a joint probability density function. One of the most-popular MCMC methods is Gibbs sampling, whose underlying Markov chain consists of a series of conditional probability density functions. In this study, since the conditional probability density functions of unknown variables are given, the Gibbs sampling approach is employed.

To generate a sample of N_{sim} sets of $(\mu, \Delta x)$ from the joint probability density function, $f_M(\mu, \Delta x \mid x_1, s_{n-1})$, we use the conditional distributions

$$f_M(\mu \mid x_1, \Delta x, s_{n-1}) = Iga\left(\mu; \alpha + \frac{(n-1)t_1 + t_2}{\nu^2}, \beta + \frac{s_{n-1} + x_1 + \Delta x}{\nu^2}\right) \tag{6.33}$$

$$f(\Delta x \mid \mu, x_1, s_{n-1}) = ga\left(\Delta x; \frac{t_2 - t_1}{\nu^2}, \frac{1}{\mu\nu^2}\right). \tag{6.34}$$

The algorithm can be summarized as in Algorithm 1. Note that $N_{burn-in}$ is the size of a burn-in period, where the obtained sample seems to be biased and is not used for further calculation. As in Algorithm 1, to generate (μ, x_1, s_{n-1}), we use the conditional distributions

$$f_M(\mu \mid \Delta x) = Iga\left(\mu; \alpha + \frac{t_2 - t_1}{\nu^2}, \beta + \frac{\Delta x}{\nu^2}\right) \tag{6.35}$$

$$f(\Delta x \mid \mu) = ga\left(\Delta x; \frac{t_2 - t_1}{\nu^2}, \frac{1}{\mu\nu^2}\right). \tag{6.36}$$

The algorithm can be summarized as in Algorithm 2.

Algorithm 1 Gibbs sampling for $(\mu, \Delta x)$ given (x_1, s_{n-1})

Obtain (x_1, s_{n-1})
Set $(\mu^0, \Delta x^0)$.
for i from 1 to $N_{burn-in} + N_{sim}$ **do**
  Generate μ^i from $f_M(\mu \mid x_1, \Delta x, s_{n-1})$
  Generate Δx^i from $f(\Delta x \mid \mu, x_1, s_{n-1})$
end for

Algorithm 2 Gibbs sampling for (μ, x_1, s_{n-1})

Set $(\mu^0, x_1^0, s_{n-1}^0)$.
for i from 1 to $N_{burn-in} + N_{sim}$ **do**
  Generate μ^i from $f_M(\mu \mid x_1^{i-1}, s_{n-1}^{i-1})$
  Generate x_1^i from $f_{X_1|M}(x_1 \mid \mu^i)$
  Generate s_{n-1}^i from $f_{S_{n-1}|M}(s_{n-1} \mid \mu^i)$
end for

6.4.2 Evaluation of *ENGS*

The calculation processes of *ENGS* for single- and multiple-component problems are described as computational algorithms.

Single-Component Problem

To derive *EGNS*, the expected prior and pre-posterior costs derived respectively in Equations (6.13) and (6.16) need to be numerically calculated. Algorithm 3 illustrates how to obtain the expected prior cost. Given μ from the prior distribution, $f_M(\mu)$, the probability of failure is derived using cumulative density function, $F_{X_2}(\rho_F)$. The expectation is calculated through Monte Carlo simulation (MCS).

Algorithm 4 explains the simulation process of the expected pre-posterior cost. After x_1 and s_{n-1} are generated through MCMC (Gibbs sampling), μ is simulated from a posterior distribution, $f_M(\mu \mid x_1)$. Given μ, the probability of failure is calculated. Note that N_K and N_J are the

Algorithm 3 Calculation of expected prior cost

Generate N_K of μ from $f_M(\mu)$ using MCS
for k from 1 to N_K **do**
  Calculate probability of failure, $1 - F_{X_2}(\rho_F)$, using $f_{X_2|M}(x_2 \mid \mu_k)$
end for
Calculate $\mathbb{E}_{\boldsymbol{Z}_2}\left[C(a_2, \boldsymbol{Z}_2)\right] \approx \frac{\sum_{k=1}^{N_K} C_F(1-F_{X_2}(\rho_F))}{N_K}$
Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_2}\left[C(a, \boldsymbol{Z}_2)\right]$

Algorithm 4 Calculation of expected pre-posterior cost: Single-component

Generate N_J of x_1 from $f_M(\mu, x_1)$ using Gibbs sampling
for j from 1 to N_J **do**
  Generate N_K of μ from $f_M(\mu \mid x_{1,j})$ using MCS
  for k from 1 to N_K **do**
    Calculate probability of failure, $1 - F_{\Delta X,k}(\rho_F - x_{1,j})$, using $f(\Delta x \mid \mu_k)$
    Calculate $C_F(1 - F_{\Delta X,k}(\rho_F - x_{1,j}))$
  end for
  Calculate $\mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}}\left[C(a_2, \boldsymbol{Z}_2 \mid e_1, x_{1,j}))\right] \approx \frac{\sum_{k=1}^{N_K} C_F(1-F_{\Delta X,k}(\rho_F - x_{1,j}))}{N_K}$
  Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_{1,j})\right]$
end for
Calculate $\mathbb{E}_{X_1}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_1)\right]\right] \approx \frac{\sum_{j=1}^{N_J} \min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}}\left[C(a, \boldsymbol{Z}_2|e_1, x_{1,j})\right]}{N_J}$

simulation sample sizes for inner and outer loops, respectively.

Multiple-Component Problem

Algorithm 3 is also used for multiple-component problems; the expected prior cost is the cost for a single-component multiplied by N. As described in Section 6.3.4, the expected pre-posterior costs for inspected and un-inspected components need to be separately considered. For inspected components, the algorithm used to numerically calculate the expected pre-posterior cost is as in Algorithm 5. The algorithm for un-inspected components has the same logic, but a slight difference exists in calculating the probability of failure in the internal loop of the algorithm, as follows in Algorithm 6.

Algorithm 5 Calculation of expected pre-posterior cost: Multiple-component (inspected)

Generate N_J sets of vector, (x_1, s_{n-1}), from $f_M(\mu, x_1, s_{n-1})$ using Gibbs sampling
for j from 1 to N_J **do**
  Generate N_K of μ from $f_M(\mu \mid x_{1,j}, s_{n-1,j})$ using MCS
  for k from 1 to N_K **do**
    Calculate probability of failure, $1 - F_{\Delta X,k}(\rho_F - x_{1,j})$, using $f(\Delta x \mid \mu_k)$
    Calculate an expected cost of taking a_2, $C_F(1 - F_{\Delta X,k}(\rho_F - x_{1,j}))$
  end for
  Calculate $\mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C_F(1 - F_{\Delta X}(\rho_F - x_{1,j})) \mid e_1, n, x_{1,j}, s_{n-1,j})\right] \approx \frac{\sum_{k=1}^{N_K} C_F(1-F_{\Delta X,k}(\rho_F - x_{1,j}))}{N_K}$
  Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_{1,j}, s_{n-1,j})\right]$
end for
Calculate $\mathbb{E}_{X_1,S_{n-1}}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_1, s_{n-1})\right]\right] \approx \frac{\sum_{j=1}^{N_J} \min_a \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a,\boldsymbol{Z}_3|e_1,n,x_{1,j},s_{n-1,j})\right]}{N_J}$

Algorithm 6 Calculation of expected pre-posterior cost: Multiple components (un-inspected)

Generate N_J of $s_n = x_1 + s_{n-1}$, from $f_M(\mu, x_1, s_{n-1})$ using Gibbs sampling
for j from 1 to N_J **do**
  Generate N_K of μ from $f_M(\mu \mid s_{n,j})$ using MCS
  for k from 1 to N_K **do**
    Calculate probability of failure, $1 - F_{X_2,k}(\rho_F)$, using $f_{X_2|M}(x_2 \mid \mu_k)$
    Calculate an expected cost of taking a_2, $C_F(1 - F_{X_2,k}(\rho_F))$
  end for
  Calculate $\mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C_F(1 - F_{X_2,k}(\rho_F))\right] \approx \frac{\sum_{k=1}^{N_K} C_F(1-F_{X_2,k}(\rho_F))}{N_K}$
  Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_{n,j})\right]$
end for
Calculate $\mathbb{E}_{S_n}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, s_n)\right]\right] \approx \frac{\sum_{j=1}^{N_J} \min_a \mathbb{E}_{\boldsymbol{Z}_3|s_n}\left[C(a,\boldsymbol{Z}_3|e_2,n,s_{n,j})\right]}{N_J}$

6.4.3 Approximate Methods for *ENGS* Evaluation

To reduce computational cost, we discretize the continuous random variables and adopt two methods: memoization and probability mass function (PMF) method.

Method 1: Memoization Technique

Memoization is used to speed up computational calculations by storing the results of previous calculations and returning them when the same inputs occur again. First, the space of the observation result, X_1, is discretized. Here, $x_{1,j}^{disc}$ is a discretized value and represents x_1 around it. Then, once a reasonable range of possible x_1 has been determined, the variation of the input of x_1 becomes finite. Due to the limited number of inputs, the memorization technique speeds up the algorithm. Algorithm 7 is for simulating the expected pre-posterior cost with memoization technique.

For a single-component problem, Algorithm 4 is modified to become Algorithm 7. Similarly, memorization is applied to Algorithms 5 and 6 for inspected and un-inspected components in a multiple-component problem, as shown in Algorithms 8 and 9, respectively.

Method 2: Probability Mass Function

The basic idea of the method introduced in this section is simple and common. Operators just need to use PMFs instead of PDFs, although this idea has a specific name, "PMF method," in this thesis for convenience. This method is meaningful only if the PDF can be analytically derived, although the algorithms with MCS or MCMC are possible for any situations. Thus, by combining the gamma process model for the degradation process with its conjugate prior distribution, inverse gamma, for a parameter, μ, we can derive analytical joint and marginal distributions of random variables without using MCS or MCMC. By the definition of probability density, the PMF can be derived from PDF as follows:

$$f_{X^{disc}}(x_j^{disc}) = \int_{x_j^{disc}-\frac{w}{2}}^{x_j^{disc}+\frac{w}{2}} f_X(x)dx, \tag{6.37}$$

Algorithm 7 Memoization method: Single-component

Discretize the field of X_1
Generate N_J of x_1 from $f_M(\mu, x_1)$ using Gibbs sampling
for j from 1 to N_J **do**
  Obtain the closest discretized variable, $x_{1,j}^{disc}$
  if $x_{1,j}^{disc}$ has never selected before **then**
    Generate N_K of μ from $f_M(\mu \mid x_{1,j}^{disc})$ using MCS
    for k from 1 to N_K **do**
      Calculate probability of failure, $1 - F_{\Delta X,k}(\rho_F - x_{1,j}^{disc})$, using $f(\Delta x \mid \mu_k)$
      Calculate an expected cost of taking a_2, $C_F(1 - F_{\Delta X,k}(\rho_F - x_{1,j}^{disc}))$
    end for
    Calculate $\mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}^{disc}}\left[C_F(1 - F_{\Delta X}(\rho_F - x_{1,j}^{disc}))\right] \approx \frac{\sum_{k=1}^{N_K} C_F(1-F_{\Delta X,k}(\rho_F - x_{1,j}^{disc}))}{N_K}$
    Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}^{disc}}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_{1,j}^{disc}\right]$
  else
    Obtain $\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}^{disc}}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_{1,j}^{disc}\right]$ from previous calculation
  end if
end for
Calculate $\mathbb{E}_{X_1}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_1}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_1)\right]\right] \approx \frac{\sum_{j=1}^{N_J} \min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}^{disc}}\left[C(a, \boldsymbol{Z}_2|e_1, x_{1,j})\right]}{N_J}$

where $f_{X^{disc}}(x^{disc})$ and $f_X(x)$ denote the PMF and PDF of X, respectively; w is the interval of the discretized value of X. When w is small enough, the PMF can be approximated as follows:

$$f_{X^{disc}}(x_j^{disc}) \approx w f_X(x_j^{disc}). \tag{6.38}$$

If the discretized region has upper and lower bounds, the PMF needs to be normalized as

$$f_{X^{disc}}(x_j^{disc}) \approx \frac{f_X(x_j^{disc})}{\sum_{j \in J} f_X(x_j^{disc})}, \tag{6.39}$$

Algorithm 8 Memoization method: Multiple-component (inspected)

Discretize the two dimensional field, (X_1, S_{n-1})
Generate N_J sets of (x_1, s_{n-1}) from $f_M(\mu, x_1, s_{n-1})$ using Gibbs sampling
for j from 1 to N_J **do**
  Obtain the closest pair of discretized variables, $(x_{1,j}^{disc}, s_{n-1,j}^{disc})$
  if $(x_{1,j}^{disc}, s_{n-1,j}^{disc})$ has never selected before **then**
    Generate N_K of μ from $f_M(\mu \mid x_{1,j}^{disc}, s_{n-1,j}^{disc})$ using MCS
    for k from 1 to N_K **do**
      Calculate probability of failure, $1 - F_{\Delta X,k}(\rho_F - x_{1,j}^{disc})$, using $f(\Delta x \mid \mu_k)$
      Calculate an expected cost of taking a_2, $C_F(1 - F_{\Delta X,k}(\rho_F - x_{1,j}^{disc}))$
    end for
    Calculate $\mathbb{E}_{\boldsymbol{Z}_3|x_{1,j}^{disc}, s_{n-1,j}^{disc}}\left[C_F(1 - F_{\Delta X}(\rho_F - x_{1,j}^{disc}))\right] \approx \frac{\sum_{k=1}^{N_K} C_F(1-F_{\Delta X,k}(\rho_F - x_{1,j}^{disc}))}{N_K}$
    Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_3|x_{1,j}^{disc}, s_{n-1,j}^{disc}}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_{1,j}^{disc}, s_{n-1,j}^{disc}\right]$
  else
    Obtain $\min_a \mathbb{E}_{\boldsymbol{Z}_3|x_{1,j}^{disc}, s_{n-1,j}^{disc}}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_{1,j}^{disc}, s_{n-1,j}^{disc}\right]$ from previous calculation
  end if
end for
Calculate $\mathbb{E}_{X_1, S_{n-1}}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_3|x_1, s_{n-1}}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_1, s_{n-1})\right]\right] \approx \frac{\sum_{j=1}^{N_J} \min_a \mathbb{E}_{\boldsymbol{Z}_3|x_{1,j}^{disc}, s_{n-1,j}^{disc}}\left[C(a, \boldsymbol{Z}_3|e_1, n, x_{1,j}^{disc}, s_{n-1,j}^{disc})\right]}{N_J}$

Algorithm 9 Memoization method: Multiple-component (un-inspected)

Discretize the field of S_n
Generate N_J of $s_n = x_1 + s_{n-1}$ from $f_M(\mu, x_1, s_{n-1})$ using Gibbs sampling
for j from 1 to N_J **do**
  Obtain the closest discretized variable, $s_{n,j}^{disc}$
  if $s_{n,j}^{disc}$ has never selected before **then**
    Generate N_K of μ from $f_M(\mu \mid s_{n,j}^{disc})$ using MCS
    for k from 1 to N_K **do**
      Calculate probability of failure, $1 - F_{X_2,k}(\rho_F)$, using $f(\Delta x \mid \mu_k)$
      Calculate an expected cost of taking a_2, $C_F(1 - F_{X_2,k}(\rho_F))$
    end for
    Calculate $\mathbb{E}_{\boldsymbol{Z}_3 \mid s_{n,j}^{disc}} \left[C_F(1 - F_{X_2,k}(\rho_F))\right] \approx \frac{\sum_{k=1}^{N_K} C_F(1-F_{X_2,k}(\rho_F))}{N_K}$
    Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_3 \mid s_{n,j}^{disc}} \left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_{n,j}^{disc}\right]$
  else
    Obtain $\min_a \mathbb{E}_{\boldsymbol{Z}_3 \mid s_{n,j}^{disc}} \left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_{n,j}^{disc}\right]$ from previous calculation
  end if
end for
Calculate $\mathbb{E}_{S_n} \left[\min_a \mathbb{E}_{M \mid s_n} \left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_n)\right]\right] \approx \frac{\sum_{j=1}^{N_J} \min_a \mathbb{E}_{\boldsymbol{Z}_3 \mid s_{n,j}^{disc}} \left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_{n,j}^{disc})\right]}{N_J}$

Algorithm 10 PMF method: Single-component

Discretize the field of X_1
for each $x_{1,j}^{disc}$ in J, **do**
  Calculate $f_{X^{disc}}(x_{1,j}^{disc})$
  Discretize the field of X_2
  Calculate an expected probability of failure, $\mathbb{E}_{M|x_{1,j}^{disc}}\left[1 - F_{\Delta X}(\rho_F - x_{1,j}^{disc})\right]$,
    using $f(\Delta x \mid x_{1,j})$
  Calculate an expected cost of taking a_2, $C_F\mathbb{E}_{M|x_{1,j}^{disc}}\left[1 - F_{\Delta X}(\rho_F - x_{1,j}^{disc})\right]$
  Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}^{disc}}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_{1,j}^{disc}\right]$
end for
Calculate $\mathbb{E}_{X_1}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_2|x_1}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_1)\right]\right] \approx \sum_{j\in J} \min_a \mathbb{E}_{\boldsymbol{Z}_2|x_{1,j}^{disc}}\left[C(a, \boldsymbol{Z}_2 \mid e_1, x_{1,j})\right] f_{X^{disc}}(x_{1,j}^{disc})$

where J is the set of all numbers for descretized X and is a finite integer. With the subset of J, L, which includes all j that satisfy $x_j^{disc} < \rho_F$, the cumulative mass function is derived as follows:

$$F_{X^{disc}}(\rho) \approx \frac{\sum_{j\in L} f_{X^{disc}}(x_j^{disc})}{\sum_{j\in J} f_X(x_j^{disc})} \tag{6.40}$$

Based on the approximated PMF method, Algorithms 4, 5, and 6 are modified as Algorithms 10, 11, and 12, respectively.

6.4.4 Efficiency of Algorithms

To illustrate the efficiency of the algorithms, 100 of *ENGS*s are simulated with each combination of the simulation sample sizes of outer and inter loops, N_J and N_K, respectively. Oakley et al. (2010) show that the simulation sample size of an outer loop largely contributes to reducing the standard deviation in simulation results, although the sample size of an inner loop changes the results somewhat. We confirm the influence of each simulation sample size and compare the results of original and memoization methods with large N_J and small N_K. For this process, we use Intel Core(TM)2 Duo with CPU 2.80 GHz and RAM 4.00 GB.

Algorithm 11 PMF method: Multiple-component (inspected)

Discretize the two dimensional field, (X_1, S_{n-1})
for j in J **do**
  for k in K **do**
    Calculate $f_{X^{disc}, S_{n-1}^{disc}}(x_{1,j}^{disc}, s_{n-1,k})$
    Calculate an expected probability of failure, $\mathbb{E}_{M|x_{1,j}^{disc}, s_{n-1,k}^{disc}}\left[1 - F_{\Delta X,k}(\rho_F - x_{1,j}^{disc})\right]$, using $f(\Delta x \mid x_{1,j}^{disc}, s_{n-1,k})$
    Calculate an expected cost of taking a_2, $C_F \mathbb{E}_{M|x_{1,j}^{disc}, s_{n-1,k}^{disc}}\left[1 - F_{\Delta X,k}(\rho_F - x_{1,j}^{disc})\right]$
    Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_3|x_{1,j}^{disc}, s_{n-1,k}^{disc}}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_{1,j}^{disc}, s_{n-1,j}^{disc}\right]$
  end for
end for
Calculate $\mathbb{E}_{X_1, S_{n-1}}\left[\min_a \mathbb{E}_{\boldsymbol{Z}_3|x_1, s_{n-1}}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_1, s_{n-1})\right]\right]$
$\approx \sum_{j \in J} \sum_{k \in K} \min_a \mathbb{E}_{\boldsymbol{Z}_3|x_{1,j}^{disc}, s_{n-1,j}^{disc}}\left[C(a, \boldsymbol{Z}_3 \mid e_1, n, x_{1,j}^{disc}, s_{n-1,k}^{disc})\right] f_{X^{disc}, S_{n-1}^{disc}}(x_{1,j}^{disc}, s_{n-1,k})$

Tables 6.2 and 6.3 summarize the results of the simulations; an average, a standard deviation, and a simulation time (minute) of each simulation setting are compared to one another. Table 6.2 demonstrates the same conclusion about the efficient balance of N_J and N_K. Increasing N_J reduces the standard deviation more than that of N_K. This result consistent with the insights from the study by Oakley et al. (2010).

Table 6.3 indicates that the PMF method is more accurate and computationally more efficient than the other two methods, although, as a Monte Carlo tequnique, the memoization method is more efficient than the original method, with almost the same average and standard deviations. For illustrative purpose, the simulation results of the two approximation methods are compared for n from 1 to 50. Figure 6.3 indicates that the results from the two methods match well, although the plots with the memoization method has errors obtained through MCSs.

Algorithm 12 PMF method: Multiple-component (un-inspected)

Discretize the field of S_n
for k in K **do**
  Calculate $f_{S_n^{disc}}(s_{n,k})$
  Calculate an expected probability of failure, $\mathbb{E}_{M|s_{n,k}^{disc}}[1 - F_{X_2,k}(\rho_F)]$, using $f_{X_2|S_n}(x_2 \mid s_{n,k})$
  Calculate an expected cost of taking a_2, $C_F \mathbb{E}_{M|s_{n,k}^{disc}}[1 - F_{X_2,k}(\rho_F)]$
  Determine $\min_a \mathbb{E}_{\boldsymbol{Z}_3|s_{n,k}^{disc}}\left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_{n,k}^{disc}\right]$
end for
Calculate $\mathbb{E}_{S_n}\left[\min_a \mathbb{E}_{M|s_n}[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_n)]\right]$
  $\approx \sum_{k \in K} \min_a \mathbb{E}_{\boldsymbol{Z}_3|s_{n,k}^{disc}}\left[C(a, \boldsymbol{Z}_3 \mid e_2, n, s_{n,k}^{disc})\right] f_{S_n^{disc}}(s_{n,k})$

6.5 Numerical Example

We demonstrate the VoI-based approach using an example in which we impose several initial settings: $N = 100$, $C_F = 100$, $C_P = 10$, $C_I = 1$, $t_1 = 25$, $t_2 = 30$, $1/\nu^2 = 9$, and $\rho_F = 3.0$. These settings represent the maintenance problem of feeder channels in the CANDU 600. Based on observed data from the feeder channels, a parameter uncertainty for μ is set as a random variable, whose prior distribution for μ is given as $Iga\,(\mu; 1102, 97.84)$. With the PMF method, Equation (2.6) is numerically solved.

6.5.1 Single-Component Problem

For the given single-component problem, the *ENGS* is calculated at 1.51; the operators should inspect the component to make better maintenance decisions. The expected prior and pre-posterior costs are 3.65 and 2.14, respectively. This result indicates that if the inspection cost, C_I, is less than 2.51, the optimal decision is "inspect the component."

Figure 6.4 shows the *ENGS* for each possible combination of C_F and C_P, and indicates that the *ENGS* is almost proportional to C_P/C_F; as C_F increases or C_P decreases, operators are

Table 6.2: Computational efficiency with different outer and inner simulation sample sizes

Method	Original method		
Outer sample size (N_J)	100	1,000	1,000
Inner sample size (N_K)	1,000	100	1,000
Average	18.7	12.2	13.0
Standard deviation	49.1	16.8	19.4
Time (minute)	26.4	26.3	60.2

Table 6.3: Computational efficiency with the original and two approximated methods

Method	Original method			Memoization method			PMF method
Outer sample size (N_J)	1,000	10,000	100,000	1,000	10,000	100,000	N.A.
Inner sample size (N_K)	1,000	100	100	1,000	100	100	N.A.
Average	13.0	13.1	12.9	12.4	13.2	12.9	12.8
Standard deviation	19.4	6.0	1.8	18.0	5.5	1.8	0
Time (minute)	60.2	60.2	411.9	36.6	21.3	95.5	9.2

expected to obtain higher *ENGS*. The *ENGS* has its peak when C_P/C_F is around 5% and drops off as C_P/C_F becomes smaller than that peak value. This tendency indicates an important characteristic of the VoI; VoI rises more when additional information has a high potential to reverse decisions originally based on the initial condition and prediction model. Since $\mathbb{P}[X_2 > \rho_F]$ with the initial settings of 0.037, when C_P/C_F is around that value, operators cannot confidently make their decision relying only on their current information.

6.5.2 Multiple-Component Problem

The optimal sample size is numerically obtained for a multiple-component inspection problem. We use the same initial settings. Without inspection, the optimal action for the prior analysis is a_2, "do-not-replace." For each sample size, the cost for an inspected/un-inspected component is calculated, and consequently, the *ENGS* is derived, as shown in Figure 6.5. The optimal sample size is 100.

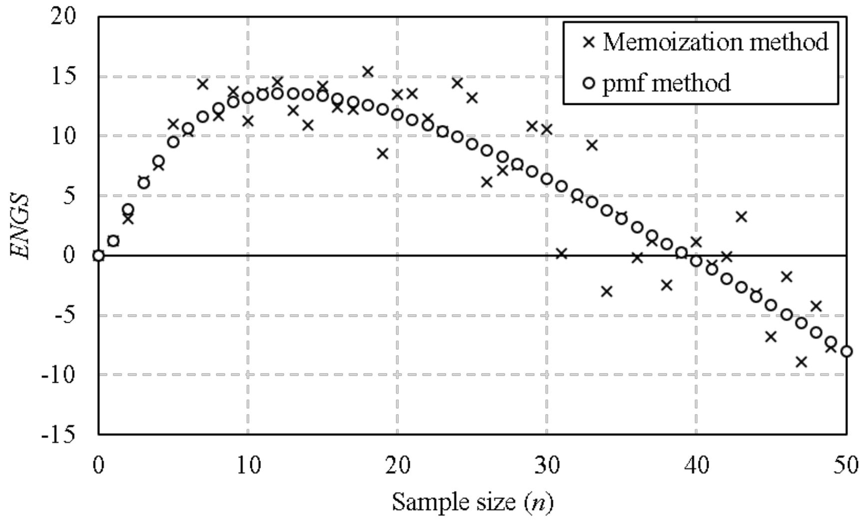

Figure 6.3: *ENGS* of the PMF and memoization ($N_J = 100,000$ and $N_K = 100$) methods for each sample size ($C_I = 3$, $C_P = 10$, and $C_F = 100$)

Quantitative Classification of Optimal Sample Sizes

Figure 6.5 illustrates the sources of the *ENGS* for each sample size. As a sample size becomes larger, the contribution of inspected components surprisingly tends to linearly increase because the expected optimal cost for an inspected component, $\bar{C}^o_{insp}$, is almost constant with respect to n. This result indicates that the VoI from inspected components is obtained mostly by reducing the temporal uncertainty instead of the parameter uncertainty. On the other hand, the contribution of un-inspected components, obtained by reducing parameter uncertainty, has its peak around $n = 20$.

The extent of the constant characteristics of $\bar{C}^o_{insp}$ is examined for different amounts of prior information, represented by the variance of the prior distribution. If a sampling inspection greatly reduces $\bar{C}^o_{insp}$ through parameter updating, $\bar{C}^o_{insp}$ would no longer be constant with respect to sample size. Figure 6.6 shows that the constant characteristics are universal. Although the

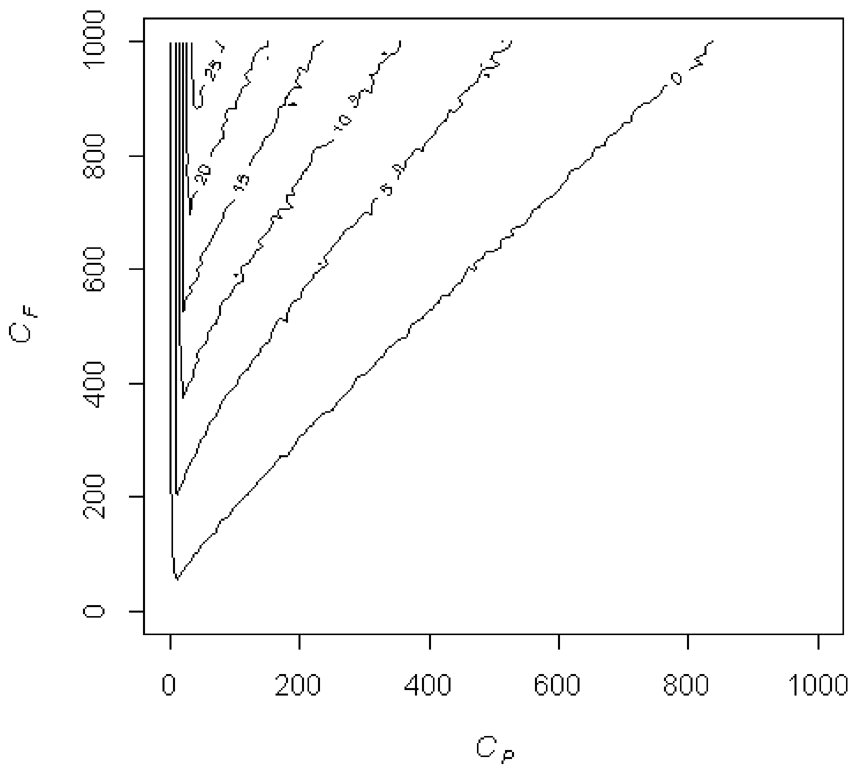

Figure 6.4: *ENGS* for each possible combination of C_F and C_P for a single-component problem with parameter uncertainty ($C_I = 1$)

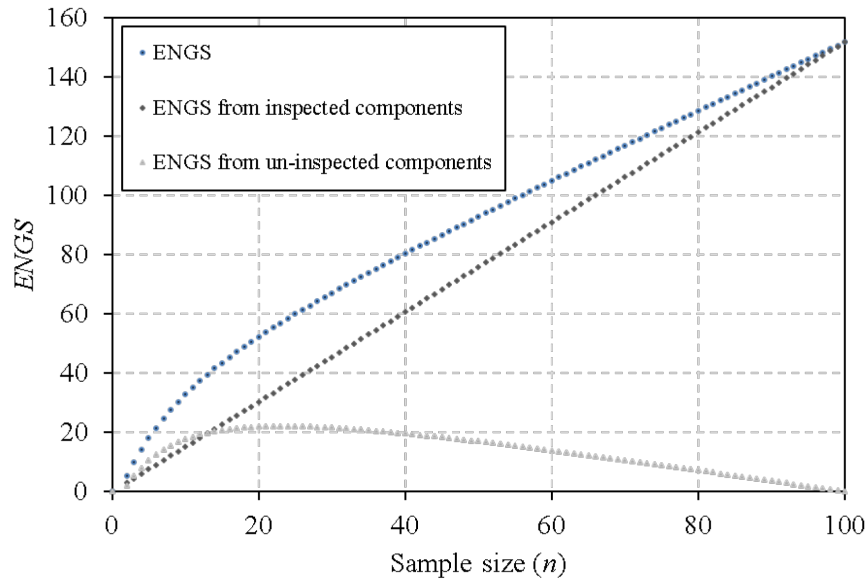

Figure 6.5: Contributions of inspected and un-inspected components for *ENGS* ($C_I = 1$, $C_P = 10$, and $C_F = 100$)

expected optimal cost becomes more costly as the amount of prior information decreases (the variance increases), the expected costs do not change for any sample size.

Through the previous calculations, the following approximations can be proposed:

- $\mathbb{E}_{X_1^{(i)}, S_{n-1}^{(-i)}}\left[C_{insp}^o\right] = \bar{C}_{insp}^o$ is a constant value with respect to n;
- $\frac{d\bar{C}_{un}^o(n)}{dn} > 0$;
- $\frac{d^2\bar{C}_{un}^o(n)}{dn^2} < 0$;

where $\bar{C}_{un}^o(n) = \mathbb{E}_{S_n}\left[C_{un}^o\right]$. Then, based on Equation (6.31), the derivative of *ENGS(n)* with respect to n is obtained as follows:

$$\frac{dENGS(n)}{dn} = \bar{C}_{un}^o(n) - \bar{C}_{insp}^o - (N-n)\frac{d\bar{C}_{un}^o(n)}{dn} \tag{6.41}$$

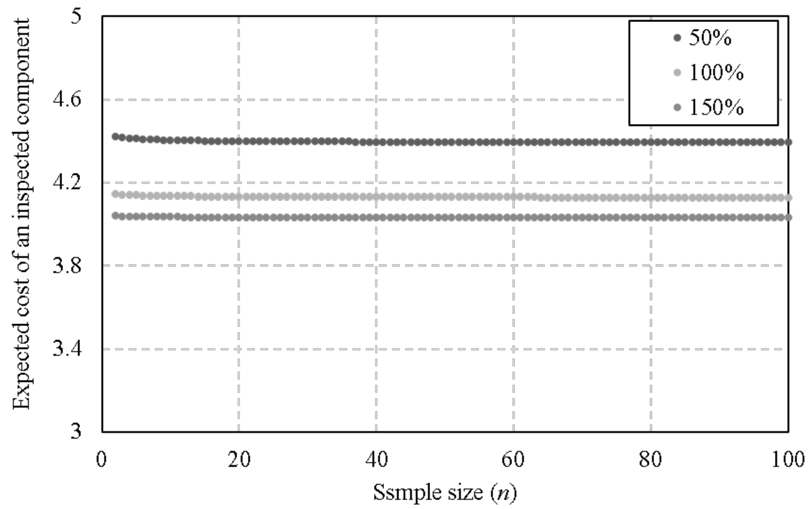

Figure 6.6: $\bar{C}^o_{insp}$ in cases with different amounts of prior information ($C_I = 1$, $C_P = 10$, and $C_F = 100$)

Thus, $n^o = N$ when

$$\lim_{n \to N-0} \frac{dENGS(n)}{dn} = \bar{C}^o_{un}(N) - \bar{C}^o_{insp} - 0 > 0$$
$$\Leftrightarrow \bar{C}^o_{insp} < \bar{C}^o_{un}(N), \tag{6.42}$$

and $n^o = 0$ when

$$\lim_{n \to +0} \frac{dENGS(n)}{dn} = \bar{C}^o_{un}(n) - \bar{C}^o_{insp} - N \lim_{n \to +0} \frac{d\bar{C}^o_{un}(n)}{dn} < 0$$
$$\Leftrightarrow \bar{C}^o_{insp} > C_{prior} - N \lim_{n \to +0} \frac{d\bar{C}^o_{un}(n)}{dn}, \tag{6.43}$$

where we assume $\bar{C}^o_{un}(0) = C_{prior}$. Since $\frac{d\bar{C}^o_{un}(n)}{dn} > 0$ for any n, if $\bar{C}^o_{insp} > C_{prior}$, then $n^o = 0$.

An insight from the results is that operators may be able to avoid a part or all of the calculation steps. Based on Equations (6.42) and (6.43),

- If $\bar{C}^o_{insp} < C^o_{un}(N)$, the optimal sample size is always $n = N$;
- If $\bar{C}^o_{insp} > C_{prior}$, the optimal sample size is always $n = 0$;
- Operators need to calculate *ENGS* for sample sizes most likely below $0.4N$, only if $C_{prior} < \bar{C}^o_{insp} < C^o_{un}(N)$.

The contribution of un-inspected components is large when operators have only a little prior information, where the contribution rapidly increases and peaks at less than $n = 10$. On the other hand, when operators have much prior information, the contribution becomes relatively small, comparable with the contribution of inspected components although its peak can be more than $0.4N$. Thus, a realistic sample size, except $n = 0$ or $n = N$, is mostly in the range of $1 \leq n < 0.4N$.

Impact of Cost Balance

The optimal sample size for each case of different inspection costs is analysed as in Figure 6.7. As the inspection cost increases, the *ENGS* to be obtained through inspection is reduced. When C_I is 1 or 2, the optimal sample size is 100, which is the population size, because the expected value of to-be-obtained information is higher than the inspection cost even at $n = 100$.

When $C_I \geq 3$, full-inspection is no longer the optimal choice. The optimal sample sizes are $n = 14$ for the case of $C_I = 3$ and $n = 7$ for the case of $C_I = 4$. The *ENGS* is positive until $n = 40$ or $n = 13$ for the case of $C_I = 3$ or $C_I = 4$, respectively. Figure 6.8 depicts the *ENGS* for each sample size in the case of $C_I = 3$. The width of the positive *ENGS* range, named the "beneficial sample size range," represents the flexibility of size-of-sampling inspection decisions. If the range is wide, operators can reasonably take a conservative sample size that is not optimal but is still beneficial for them. In these cases, the expected prior cost is in between the expected

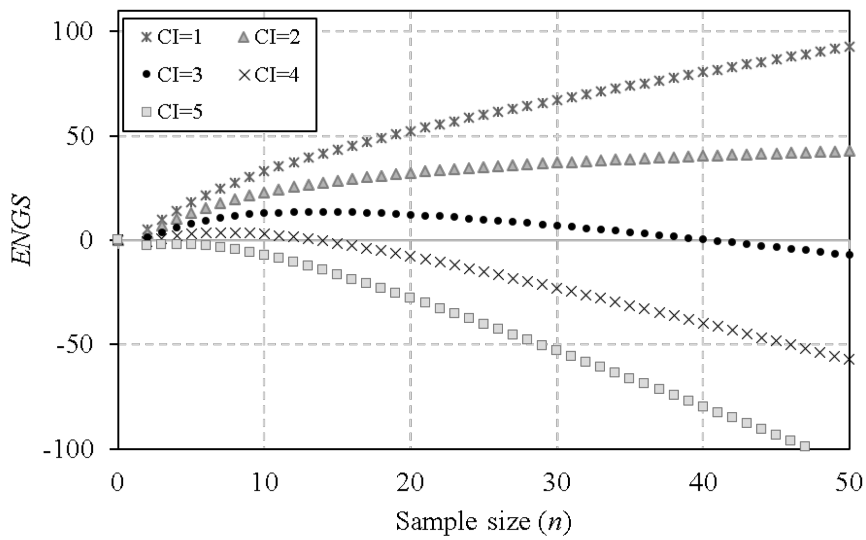

Figure 6.7: *ENGS* for different inspection costs, C_I ($C_P = 10$ and $C_F = 100$)

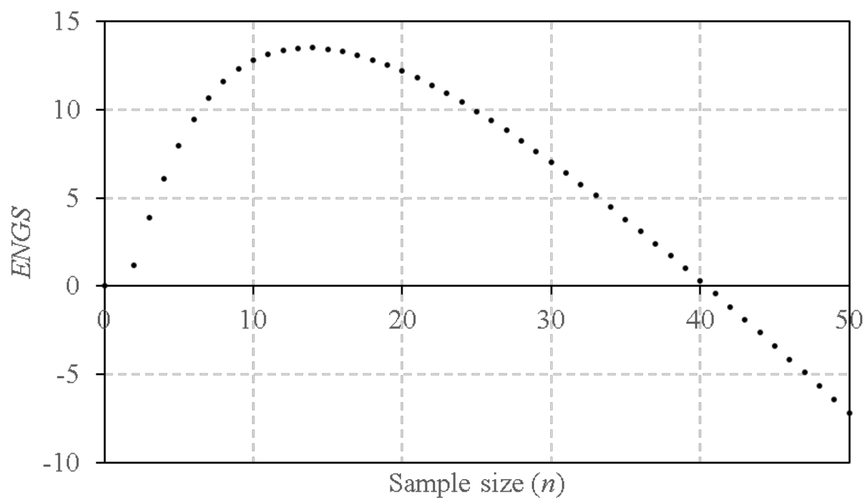

Figure 6.8: *ENGS* for the case with $C_I = 3$, $C_P = 10$, and $C_F = 100$)

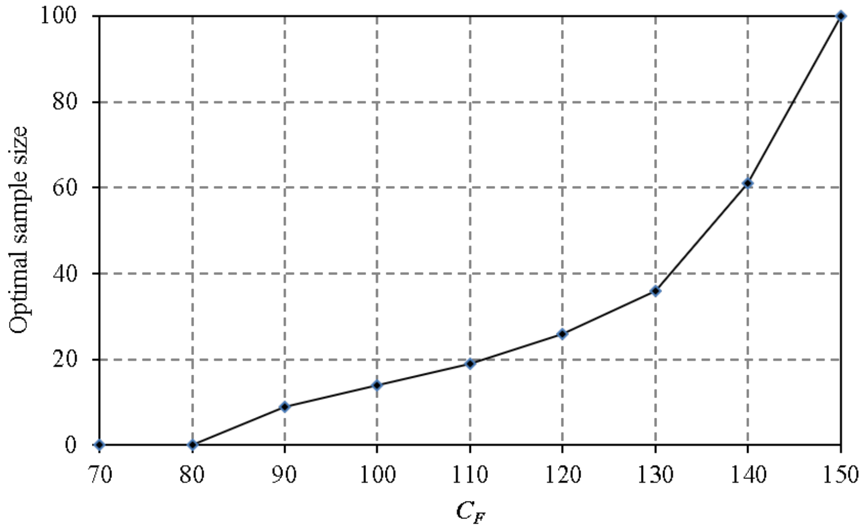

Figure 6.9: Optimal sample size for different failure costs, C_F ($C_I = 3$ and $C_P = 10$)

pre-posterior costs for inspected and un-inspected components; the benefit of reducing temporal uncertainty is not enough to compensate for the inspection cost, but the benefit of reducing parameter uncertainty contributes to making the *ENGS* positive. When n is more than four, the optimal action at the inspection stage is "do-nothing."

The influence of the failure and replacement costs, C_F and C_P, respectively, on the optimal sample size is analysed. As C_F becomes larger, the optimal sample size increases and reaches the population size (see Figure 6.9). In contrast, when C_P is changed as shown in Figure 6.10, the optimal sample size has a maximum around $C_P = 7$. These results indicate that the optimal sample size is more sensitive with C_F than C_P.

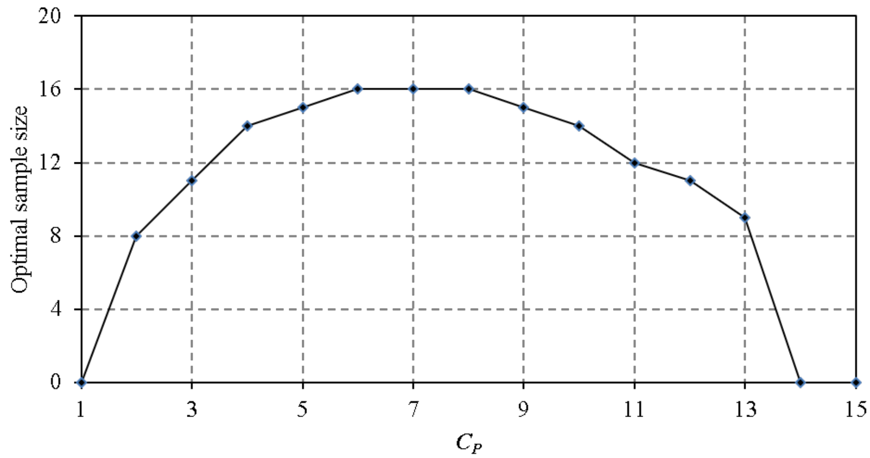

Figure 6.10: Optimal sample size for different replacement costs, C_P ($C_I = 3$ and $C_F = 100$)

Impact of Prior Information

The influence of the amount of prior information is analysed as in Figure 6.11. With the fixed ratio of α/β, α and β are changed from 10 % to 130 % of the original values, $(\alpha, \beta) = (1102, 97.8)$. As the amount of prior information increases, the optimal sample size linearly decreases and drops to zero at 130 %. This explains that a certain amount of information exists at which operators should not take into account additional information that reduces uncertainties in terms of cost-benefit analysis. At between 120 % and 130 % of the amount of the original information, the *ENGS* becomes negative over all n, although the peak of the *ENGS* contributed by un-inspected components continues to linearly decrease even after 130 %.

Impact of Inspection Timing

We have compared costs with different inspection timings. Figure 6.12 shows that the sooner operators can carry out an inspection, the smaller the *ENGS* value they will obtain; however, the

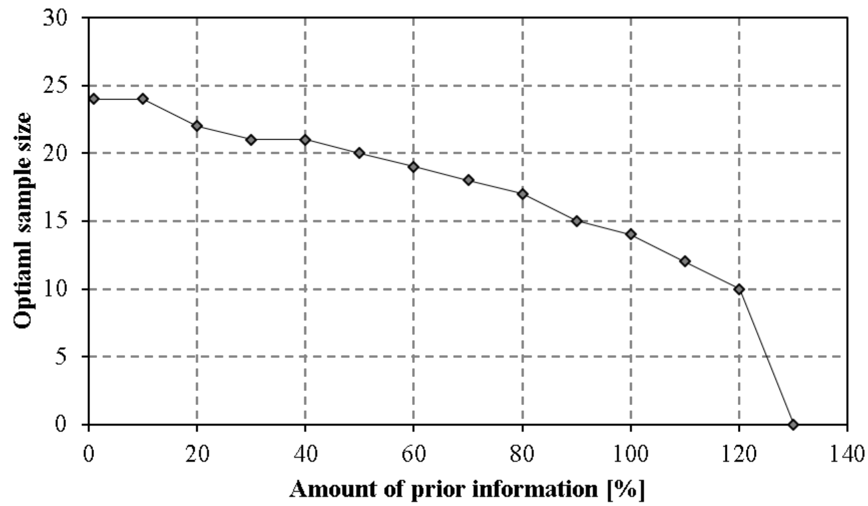

Figure 6.11: Optimal sample sizes for different prior information ($C_I = 3$, $C_P = 10$, and $C_F = 100$)

optimal sample sizes are similar. In cases of $t_1 = 25$, $t_1 = 20$, and $t_1 = 15$, the optimal sample sizes are $n = 13$, $n = 10$, and $n = 0$, respectively. When the probability of failure before t_1 is significantly low and ignorable, the later operators inspect components, the more they can reduce both the temporal and parameter uncertainties.

6.6 Summary

This chapter has developed the VoI-based sample size determination method for the maintenance decision-making problem under an assumption of dependent components through shared parameter uncertainty. Based on the model defined in Chapter 5, we have developed the model so as to deal with both temporal (aleatory) and parameter (epistemic) uncertainties in the maintenance problem. With the gamma process, we demonstrated how the observation and following updating process can be simplified in the mathematical equations. Computational

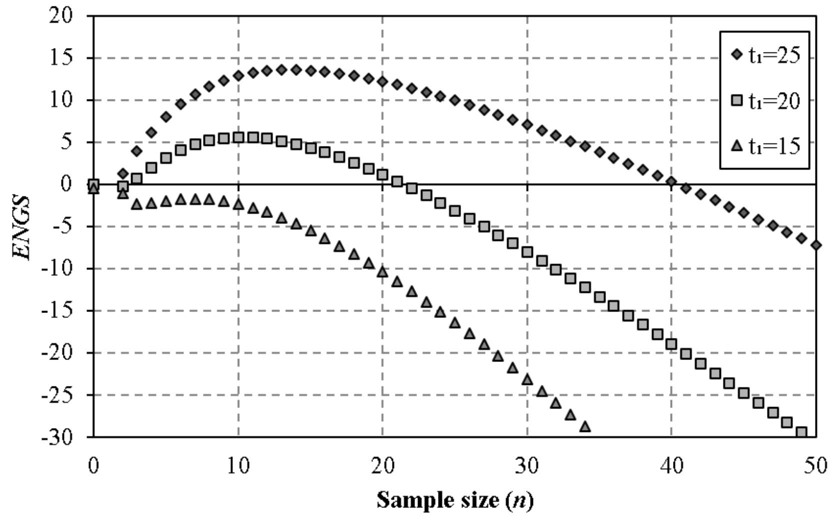

Figure 6.12: *ENGS*s with different inspection timings ($C_I = 3$, $C_P = 10$, and $C_F = 100$)

algorithms are introduced and computationally efficient algorithms have been discussed. We have demonstrated the proposed method through a numerical example. In addition to the optimal sample size, we propose a new index for sampling inspection, the "beneficial sample size range," which represents the flexibility of SSD. By changing major parameters of the model, we explored how the proposed method can apply to a variety of cases.

The major insights found for sample size determination strategies are as follows:

- The optimal sample size is sensitive against parameters of cost and prior distribution;
- The most sensitive parameter is inspection cost;
- If $\bar{C}^o_{insp} < C^o_{un}(N)$, the optimal sample size is always $n = N$;
- If $\bar{C}^o_{insp} > C_{prior}$, the optimal sample size is always $n = 0$;
- Operators generally need to calculate *ENGS* for sample sizes below $0.4N$, only if $C_{prior} < \bar{C}^o_{insp} < C^o_{un}(N)$.

Chapter 7

Two-Inspection Problem: Value of Information Analysis

Repeated inspection has been a common situation in condition-based maintenance studies. Periodic inspection, which is one of the repeated inspection policies, is widely used, for instance, Pandey et al. (2009). Under the inspection policy, a component is inspected periodically without consideration of different inspection options. The policy is simple, flexible, and an optimal solution is easy to find, but sample size determination (SSD) cannot be included in the analysis. The inspection actions need to be defined as a part of a sequential decision problem. A series of studies has developed optimization approaches for the condition-based maintenance problem including inspection actions by modelling the problem as a partially observable Markov decision process (POMDP).

Studies using POMDPs still use a fixed action interval but enable decision makers to decide an inspection action at each horizon. These studies successfully generalize the discrete case of inspection optimization in a condition-based maintenance problem for a single-component system(Papakonstantinou and Shinozuka, 2014b,c,a; Papakonstantinou and Memarzadeh, 2017). The proposed method discretizes all the factors in the problem: state of the component, inspection and maintenance actions, and observation outcomes. Schöbi and Chatzi (2016) extend the POMDP model for continuous-state problem. However, these studies focus only on a single-component system and cannot be applied for a multi-component problem. The SSD problem cannot be formulated with their approaches. Memarzadeh and Pozzi (2016) propose a method for multi-component system and evaluate the expected value of sample information (*EVSI*) for two predetermined scenarios: optimistic and pessimistic. The study offers insights into how much impact the first inspection has for the given scenarios, but it cannot explain how the first inspection affects the following inspection actions.

This chapter describes how to determine the sample size for a multiple-component system maintenance problem with multiple inspections. The study adopts dynamic programming as a basis for the method for modelling and solving the problem. First, the background and main ideas of dynamic programming are introduced. Classifying the related studies that use dynamic programming for maintenance problems, we highlight the limitations of these studies. Similar to Chapter 6, we explain how the gamma process contributes for modelling a condition-based maintenance model with multiple-inspection problems. A mathematical derivation process for the net benefit of inspection (*ENGS*) guides readers to an understanding of maintenance problems as dynamic programming problems. Numerical analysis with a real case study shows the applicability of the stated method in realistic situations.

7.1 Problem Definition

Let us assume that operators are going to choose an inspection and maintenance policy for an N-component system in operation. The situation is the same as that in the problem described in Chapter 6, except that there is now the chance of a second inspection and/or maintenance at t_2. Note that, in this chapter, t_1 and t_2 denote the first and second inspection/maintenance times,

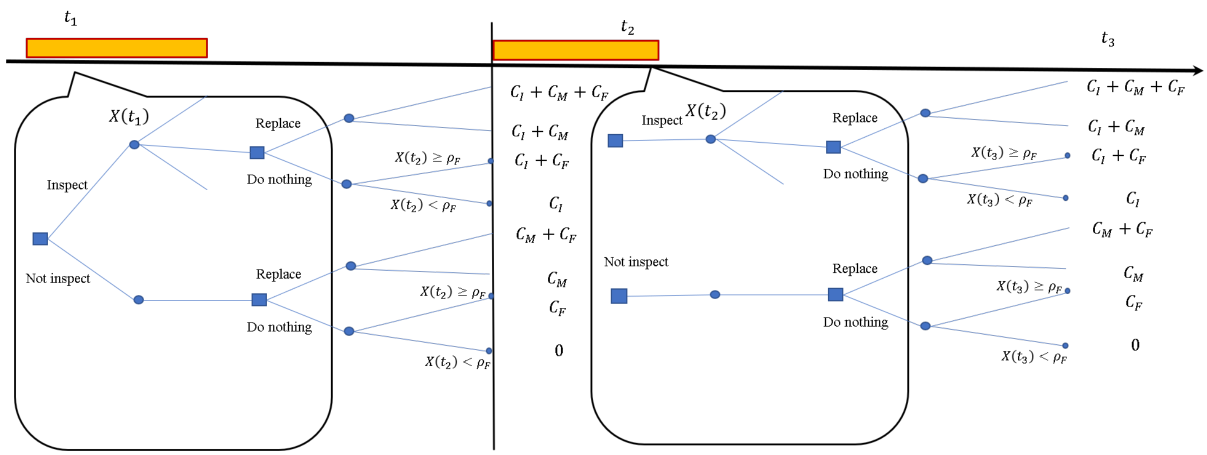

Figure 7.1: Decision tree for two-inspection problem

and t_3 is the system-decommission time.

We have two decision-making horizons, t_1 and t_2, and each horizon is divided into two stages: inspection and maintenance. The operators are at t_1 and need to take a combination of inspection and maintenance actions, which are, respectively, deciding on the size of the inspection sampling and making a decision to replace or not for each component. At the inspection stage, the inspection decision is for the whole system (all components), whereas at the maintenance stage, decisions are for individual components. Specifically, during the inspection stage, the operators determine a cost-effective scope (size) for the sampling inspection, $n^{(i)}$, and during the maintenance stage at t_i, they need to decide which (if any) components to replace and consider all components individually ("replace," $a_1^{(i)}$, or "do-nothing," $a_2^{(i)}$). As a rough sketch, the problem can be illustrated as a repeat of the problem used in Chapter 6 (see Figure 7.1).

7.2 Dynamic Programming for Condition-based Maintenance

Dynamic programming was characterized and studied by Bellman (1957) based on statistical sequential analysis (Wald, 1947). Dynamic programming describes a class of problems that can be divided into sub-problems. A decision-maker first optimizes these sub-problems and gradually extends his/her focusing problem toward the whole problem. For example, if the problem involves a multiple-horizon (times) decision-making problem, the analysis starts from the last horizon, and then the focusing horizon moves backward, using the results obtained for later horizons. The optimized value at a certain horizon, $V(x_t)$, can be simplified as a Bellman equation:

$$V(x_t) = \max_{a_t \in A}(R(x_t, a_t) + \gamma V(T(x_t, a_t)), \tag{7.1}$$

where x_t and a_t are a state of a system and a decision maker's action at t, respectively; $R(x_t, a_t)$ is the reward obtained at t; $T(x_t, a_t)$ is a transition for the next horizon given x_t and a_t; and γ denotes a constant discount rate with which we estimate a net present value. When $T(x_t, a_t)$ represents probabilistic transition, the decision maker needs to evaluate an expected future value as follows:

$$V(x_t) = \max_{a_t \in A}\left(R(x_t, a_t) + \gamma \int_{x_{t+1} \in X} V(x_{t+1}) \mathbb{P}\left[x_{t+1} \mid x_t, a_t\right] dx_{t+1}\right). \tag{7.2}$$

Note that summation is used instead of integration if the state-space of a system is discrete. Although dynamic programming has been developed and applied primarily in computer science, several studies have adopted it for maintenance problems. Since general dynamic programming is computationally expensive, its simplified forms, Markov decision process (MDP) and POMDP, have been widely adopted.

7.2.1 Markov Decision Process

MDP is defined as a mixture of a Markov process and a decision-making process. In a Markov process, a future condition depends only on a current state. In other words, it is memoryless; the history of previous conditions does not affect its stochastic transition to the future condition. A gamma process is one example of a continuous-state Markov process. By modelling a component's degradation as a Markov process, maintenance decision problems have been defined as MDPs. For simple formulations, a discrete-state Markov process is adopted in most studies. For example, Ahmadi (2016) presents a condition-based maintenance model based on an MDP for a single-component system. Nguyen et al. (2013) apply a MDP model for analysing optimal maintenance, including the influence of spare parts inventory.

7.2.2 Partially Observable Markov Decision Process

To include the inspection decision problem in a condition-based maintenance optimization, POMDP has been developed based on MDP. Instead of an actual state of a component, the belief of a decision maker is updated based on observation results. POMDP has been developed in reinforced learning, an area of machine learning, but has gained attention in maintenance decision problems in recent years. In the context of structural health monitoring, inspection actions are combined with maintenance actions, and optimal policies for each possible initial belief are determined. For example, Faddoul et al. (2011) analyse a two sequence inspection problem through which they explain the approach in a POMDP. Zhang and Revie (2017) develop a partially observable semi-Markov decision process (POSMDP) for continuous-time imperfect inspection problems.

7.3 Single-Component Problem

This chapter builds the model for a two-inspection problem in which the operators' primary focus is determining the sample size for the first inspection, $n^{(1)}$. In that context, the prior analysis means the optimization of the expected cost without the first sampling inspection and "with" the second sampling inspection. In short, the prior analysis consists of the pre-posterior analysis described in Chapter 6. Based on the problem in Chapter 6, the single-component problem is first stated for illustrative reasons, and the multiple-component model is next introduced.

The assumptions we set in this stated problem, which have already been introduced in Chapters 5 and 6, are re-written as follows:

- Degradation of components follows the gamma process model;
- Components are statistically independent of one another;
- The probability of a replaced component failing before t_3 is negligible;
- Every component survives until t_1 ($x_1 < \rho_F$).

7.3.1 Random Variables

The vector of the random variables, $\boldsymbol{Z}_4 = (M, X_1, X_2, X_3)$, is the source of uncertainties. The *ENGS* is obtained by reducing these uncertainties through observing $\boldsymbol{Z}_4$. The probability density functions (PDFs) and cumulative density functions (CDFs) of X_1 and X_2 are the same as those shown in Section 6.2.2, although several notations need to be modified from Δt and Δx to Δt_{12} and Δx_{12}. The condition at t_3, X_3, is estimated based on the PDF, as follows:

$$f_{X_3|M,X_1}(x_3 \mid \mu, x_1) = g(x_3 - x_1), \tag{7.3}$$

$$f_{X_3|M,X_2}(x_3 \mid \mu, x_2) = g(x_3 - x_2), \tag{7.4}$$

The increments of the degradation level over $\Delta t_{13} = t_3 - t_1$, ΔX_{13}, and $\Delta t_{23} = t_3 - t_2$, ΔX_{23}, have the following PDFs

$$\begin{aligned} f_{\Delta X_{13}|M,X_1}(\Delta x_{13} \mid \mu, x_1) &= g(\Delta x_{13}), &(7.5)\\ f_{\Delta X_{23}|M,X_2}(\Delta x_{23} \mid \mu, x_2) &= g(\Delta x_{23}). &(7.6) \end{aligned}$$

The joint distribution of M and X_3 given $X_1 = x_1$ is

$$\begin{aligned} & f_{M,X_3|X_1}(\mu, x_3 \mid x_1) \\ = \; & f_{X_3|M,X_1}(x_3 \mid \mu, x_1) f_M(\mu \mid x_1) \\ = \; & \frac{\left(\frac{x_3-x_1}{\nu^2}\right)^{\frac{\Delta t_{13}}{\nu^2}} (\beta + \frac{x_1}{\nu^2})^{\alpha+\frac{t_1}{\nu^2}}}{(x_3 - x_1)\Gamma(\frac{\Delta t_{13}}{\nu^2})\Gamma(\alpha + \frac{t_1}{\nu^2})} \left(\frac{1}{\mu}\right)^{\alpha+\frac{t_1+\Delta t_{13}}{\nu^2}+1} \exp\left(-\frac{1}{\mu}\left(\beta + \frac{x_1 + (x_3 - x_1)}{\nu^2}\right)\right), \quad (7.7) \end{aligned}$$

The marginal distribution of X_3 given $X_1 = x_1$ is

$$\begin{aligned} f_{X_3|X_1}(x_3 \mid x_1) &= \int_0^\infty f_{M,X_3|X_1}(\mu, x_3 \mid x_1) d\mu \\ &= \frac{\left(\frac{x_3-x_1}{\nu^2}\right)^{\frac{\Delta t_{13}}{\nu^2}} (\beta + \frac{x_1}{\nu^2})^{\alpha+\frac{t_1}{\nu^2}} \Gamma(\alpha + \frac{t_1+\Delta t_{13}}{\nu^2})}{(x_3 - x_1)(\beta + \frac{x_1+(x_3-x_1)}{\nu^2})^{\alpha+\frac{t_1+\Delta t_{13}}{\nu^2}} \Gamma(\frac{\Delta t_{13}}{\nu^2})\Gamma(\alpha + \frac{t_1}{\nu^2})}, \quad (7.8) \end{aligned}$$

The joint distribution of M and X_3 given $X_2 = x_2$ is

$$\begin{aligned} & f_{M,X_3|X_2}(\mu, x_3 \mid x_2) \\ = \; & f_{X_3|M,X_2}(x_3 \mid \mu, x_2) f_M(\mu \mid x_2) \\ = \; & \frac{\left(\frac{x_3-x_2}{\nu^2}\right)^{\frac{\Delta t_{23}}{\nu^2}} (\beta + \frac{x_2}{\nu^2})^{\alpha+\frac{t_2}{\nu^2}}}{(x_3 - x_2)\Gamma(\frac{\Delta t_{23}}{\nu^2})\Gamma(\alpha + \frac{t_2}{\nu^2})} \left(\frac{1}{\mu}\right)^{\alpha+\frac{t_2+\Delta t_{23}}{\nu^2}+1} \exp\left(-\frac{1}{\mu}\left(\beta + \frac{x_2 + (x_3 - x_2)}{\nu^2}\right)\right), \quad (7.9) \end{aligned}$$

The marginal distribution of X_3 given $X_2 = x_2$ is

$$\begin{aligned} f_{X_3|X_2}(x_3 \mid x_2) &= \int_0^\infty f_{M,X_3|X_2}(\mu, x_3 \mid x_2) d\mu \\ &= \frac{\left(\frac{x_3-x_2}{\nu^2}\right)^{\frac{\Delta t_{23}}{\nu^2}} (\beta + \frac{x_2}{\nu^2})^{\alpha+\frac{t_2}{\nu^2}} \Gamma(\alpha + \frac{t_2+\Delta t_{23}}{\nu^2})}{(x_3 - x_2)(\beta + \frac{x_2+(x_3-x_2)}{\nu^2})^{\alpha+\frac{t_2+\Delta t_{23}}{\nu^2}} \Gamma(\frac{\Delta t_{23}}{\nu^2})\Gamma(\alpha + \frac{t_2}{\nu^2})}, \end{aligned} \tag{7.10}$$

The CDFs of X_3 are

$$F_{X_3|M,X_1}(\rho) = \int_{x_1}^{\rho} f_{X_3|M,X_1}(x_3 \mid \mu, x_1) dx_3 = \int_0^{\rho-x_1} g(\Delta x_{13}) d\Delta x_{13}, \tag{7.11}$$

$$\begin{aligned} F_{X_3|X_1}(\rho) &= \int_{x_1}^{\rho} f_{X_3|X_1}(x_3 \mid x_1) dx_3 \\ &= \int_0^{\rho-x_1} \frac{\left(\frac{\Delta x_{13}}{\nu^2}\right)^{\frac{\Delta t_{13}}{\nu^2}} (\beta + \frac{x_1}{\nu^2})^{\alpha+\frac{t_1}{\nu^2}} \Gamma(\alpha + \frac{t_1+\Delta t_{13}}{\nu^2})}{\Delta x_{13}(\beta + \frac{x_1+\Delta x_{13}}{\nu^2})^{\alpha+\frac{t_1+\Delta t_{13}}{\nu^2}} \Gamma(\frac{\Delta t_{13}}{\nu^2})\Gamma(\alpha + \frac{t_1}{\nu^2})} d\Delta x_{13} \end{aligned} \tag{7.12}$$

$$F_{X_3|M,X_2}(\rho) = \int_{x_2}^{\rho} f_{X_3|M,X_2}(x_3 \mid \mu, x_2) dx_3 = \int_0^{\rho-x_2} g(\Delta x_{23}) d\Delta x_{23}, \tag{7.13}$$

$$\begin{aligned} F_{X_3|X_2}(\rho) &= \int_{x_2}^{\rho} f_{X_3|X_2}(x_3 \mid x_2) dx_3 \\ &= \int_0^{\rho-x_2} \frac{\left(\frac{\Delta x_{23}}{\nu^2}\right)^{\frac{\Delta t_{23}}{\nu^2}} (\beta + \frac{x_2}{\nu^2})^{\alpha+\frac{t_2}{\nu^2}} \Gamma(\alpha + \frac{t_2+\Delta t_{23}}{\nu^2})}{\Delta x_{23}(\beta + \frac{x_2+\Delta x_{23}}{\nu^2})^{\alpha+\frac{t_2+\Delta t_{23}}{\nu^2}} \Gamma(\frac{\Delta t_{23}}{\nu^2})\Gamma(\alpha + \frac{t_2}{\nu^2})} d\Delta x_{23}. \end{aligned} \tag{7.14}$$

7.3.2 Baseline Case: No inspection at t_1

The major structure of the optimization analysis without the first sampling inspection is actually the same as the structure of the pre-posterior analysis in Section 6.2, except for the shift of the time horizon and the additional maintenance decision at t_1. In both cases there is a chance to inspect once in the whole decision-making process. The expected optimal cost with no inspection at either t_1 or t_2 is calculated thus:

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_2^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] \\
= &\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_2^{(2)}, a_2^{(1)})\right] \\
= &\min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_M\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}\right). \qquad (7.15)
\end{aligned}$$

where $\boldsymbol{a}_1 = (e^{(1)}, e^{(2)}, a^{(1)}, a^{(2)})$ is the vector of the inspection and maintenance action taken at t_1 and t_2; $e_i^{(j)}$ and $a_i^{(j)}$ denote $e^{(j)} = e_i$ and $a^{(j)} = a_i$, respectively; $a^{(j)o}$ means the optimal action at t_j.

If another situation arises in which operators have an inspection result $X_2 = x_2$, the expected cost with inspection at t_2 given $X_2 = x_2$ is formulated as

$$\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4|X_2}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_1^{(2)}, a_2^{(1)}, x_2)\right] = C_I + \min_{a^{(2)}} \left(C_P, C_F \mathbb{E}_{M|X_2}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]\right). \qquad (7.16)$$

Similar to the cost in Equation (6.16), if operators do not actually have inspection results, X_2 needs to be considered as a random variable. The expected cost with respect to X_2 becomes

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_1^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] \\
= &\mathbb{E}_{X_2|X_2<\rho_F}\left[\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4|X_2}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_1^{(2)}, a_2^{(1)}, x_2)\right]\right] \\
= &\frac{\int_0^{\rho_F} \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4|X_2}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_1^{(2)}, a_2^{(1)}, x_2)\right] f_{X_2}(x_2) dx_2}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}. \qquad (7.17)
\end{aligned}$$

Note that this expected optimal cost is for the case of the component not failing before t_2, whose probability is $F_{X_2}(\rho_F)$. Based on the derived expected costs, the $ENGS^{(2)}$, which is the net benefit of inspection at t_2 (not t_1), is derived as follows:

$$\begin{aligned} ENGS^{(2)}(e^{(2)} \mid e_2^{(1)}, a_2^{(1)}) &= \mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_2^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] \\ &\quad -\mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e_1^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right]. \end{aligned} \tag{7.18}$$

Thus, operators can determine the optimal inspection action at t_2 that leads to lower expected cost than the other option as follows:

$$e^{(2)o} = \arg\max_{e^{(2)}} ENGS^{(2)}(e^{(2)} \mid e_2^{(1)}, a_2^{(1)}). \tag{7.19}$$

At the maintenance stage at t_1, the operators choose an action between a_1, replacement, and a_2, no-action. The choice is based on which option has a lower expected cost. The “prior” cost of the two-inspection problem is

$$\begin{aligned} C^{prior} &= \min_{a^{(1)}}\Big(C_P, F_{X_2}(\rho_F)\mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] \\ &\quad +C_F(1 - F_{X_2}(\rho_F))\Big), \end{aligned} \tag{7.20}$$

where $\mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right]$ is the expected cost with optimal inspection and action at t_2, $e^{(2)o}$ and $a^{(2)o}$, respectively, given $e_2^{(1)}$. The consequence of each combination of inspection and action options is summarized in the lower half ($e^{(1)} = e_2$ case) of Tables 7.1, 7.2, 7.3, and 7.4. These tables show the calculation procedures.

Table 7.1: Costs arising between t_2 and t_3 given the options at t_1 ($a^{(1)}$ and $e^{(1)}$) and $X_2 < \rho_F$

$e^{(1)}$	x_1	$a^{(1)}$	State	$e^{(2)}$	x_2	$a^{(2)}$	State	$C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid a_2^{(1)}, X_2 < \rho_F)$	Probability
e_1	✓	a_1						—	—
		a_2	Fail					—	—
			Survive	e_1	✓	a_1		$C_I + C_P$	1
						a_2	Fail	$C_I + C_F$	$\mathbb{E}_{M\mid X_2}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
							Survive	C_I	$\mathbb{E}_{M\mid X_2}\left[F_{\Delta X_{23}}(\rho_F - x_2)\right]$
				e_2	×	a_1		C_P	1
						a_2	Fail	C_F	$\frac{\mathbb{E}_{M\mid X_1}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid X_1}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}$
							Survive	0	$\frac{\mathbb{E}_{M\mid X_1}\left[F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid X_1}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}$
e_2	×	a_1						—	—
		a_2	Fail					—	—
			Survive	e_1	✓	a_1		$C_I + C_P$	1
						a_2	Fail	$C_I + C_F$	$\mathbb{E}_{M\mid X_2}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
							Survive	C_I	$\mathbb{E}_{M\mid X_2}\left[F_{\Delta X_{23}}(\rho_F - x_2)\right]$
				e_2	×	a_1		C_P	1
						a_2	Fail	C_F	$\frac{\mathbb{E}_M\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}$
							Survive	0	$\frac{\mathbb{E}_M\left[F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}$

Table 7.2: Costs arising between t_1 and t_3 given the options at t_1 ($a^{(1)}$ and $e^{(1)}$)

$e^{(1)}$	x_1	$a^{(1)}$	State	$e^{(2)}$	x_2	$a^{(2)}$	$\mathbb{E}_{\boldsymbol{Z}_4\mid X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid a_2^{(1)}, X_2 < \rho_F)\right]$
e_1	✓	a_1					—
		a_2	Fail				—
			Survive	e_1	✓	a_1	$C_I + C_P$
						a_2	$C_I + C_F\mathbb{E}_{M\mid X_2}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
				e_2	×	a_1	C_P
						a_2	$C_F\frac{\mathbb{E}_{M\mid X_1}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid X_1}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}$
e_2	×	a_1					—
		a_2	Fail				—
			Survive	e_1	✓	a_1	$C_I + C_P$
						a_2	$C_I + C_F\mathbb{E}_{M\mid X_2}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
				e_2	×	a_1	C_P
						a_2	$C_F\frac{\mathbb{E}_M\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}$

Table 7.3: Costs arising between t_1 and t_3

$e^{(1)}$	x_1	$a^{(1)}$	State	$e^{(2)}$	x_2	$\mathbb{E}_{\boldsymbol{Z}_4\mid X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right]$
e_1	✓	a_1				—
		a_2	Fail			—
			Survive	e_1	✓	$C_I + \min_{a^{(2)}}(C_P, C_F\mathbb{E}_{M\mid X_2}[1 - F_{\Delta X_{23}}(\rho_F - x_2)]$
				e_2	×	$\min_{a^{(2)}}(C_P, C_F\frac{\mathbb{E}_{M\mid X_1}[F_{\Delta X_{12}}(\rho_F-x_1)-F_{\Delta X_{13}}(\rho_F-x_1)]}{\mathbb{E}_{M\mid X_1}[F_{\Delta X_{12}}(\rho_F-x_1)]})$
e_2	×	a_1				—
		a_2	Fail			—
			Survive	e_1	✓	$C_I + \min_{a^{(2)}}(C_P, C_F\mathbb{E}_{M\mid X_2}[1 - F_{\Delta X_{23}}(\rho_F - x_2))]$
				e_2	×	$\min_{a^{(2)}}(C_P, C_F\frac{\mathbb{E}_M[F_{X_2}(\rho_F)-F_{X_3}(\rho_F)]}{\mathbb{E}_M[F_{X_2}(\rho_F)]})$

Table 7.4: Expected pre-posterior cost for each combination of options at t_1

$e^{(1)}$	x_1	$a^{(1)}$	$\mathbb{E}_{\boldsymbol{Z}_4}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o}\right]$
e_1	✓	a_1	$C_I + C_P$
		a_2	$C_I + C_F\mathbb{E}_{M\mid X_1}[1 - F_{\Delta X_{12}}(\rho_F - x_1)]$ $+\mathbb{E}_{\boldsymbol{Z}_4\mid X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right]\mathbb{E}_{M\mid X_1}[F_{\Delta X_{12}}(\rho_F - x_1)]$
e_2	×	a_1	C_P
		a_2	$C_F\mathbb{E}_M[1 - F_{X_2}(\rho_F)]$ $+\mathbb{E}_{\boldsymbol{Z}_4\mid X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_2^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right]\mathbb{E}_M[F_{X_2}(\rho_F)]$

7.3.3 Posterior Analysis: Inspection and Maintenance at t_2

Suppose that operators have obtained an outcome for a sampling inspection, $X_1 = x_1 < \rho_F$. The operators will then choose the best action according to a comparison between the expected costs with updated information about M, X_2, and X_3. Because of the assumption, a replaced component will not fail. With the same posterior distribution as in Equation (6.2), the optimal expected cost arising between t_2 and t_3, given $a^{(1)} = a_2$ and $e^{(2)} = e_2$, is calculated as

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{Z}_4\mid X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_2^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] \\
= &\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4\mid X_1, X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_2^{(2)}, a_2^{(1)}, x_1)\right] \\
= &\min_{a^{(2)}}\left(C_P, C_F\frac{\mathbb{E}_{M\mid X_1}[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)]}{\mathbb{E}_{M\mid X_1}[F_{\Delta X_{12}}(\rho_F - x_1)]}\right).
\end{aligned} \tag{7.21}$$

When $a^{(1)} = a_2$, $e^{(2)} = e_1$, and $X_2 = x_2 < \rho_F$ are given, the optimal expected cost is

$$
\begin{aligned}
& \mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_1^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] \\
= \ & \mathbb{E}_{X_2|X_2<\rho_F}\left[\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4|X_2}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_1^{(2)}, a_2^{(1)}, x_2)\right]\right] \\
= \ & C_I + \min_{a^{(2)}}\left(C_P, C_F \mathbb{E}_{X_2|X_2<\rho_F}\left[\mathbb{E}_{M|X_2}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]\right]\right). \qquad (7.22)
\end{aligned}
$$

In the two-inspection problem, the second inspection decision is included in the posterior analysis part. Based on the derived expected costs, the $ENGS^{(2)}$, which is the net benefit of inspection at t_2, is derived as follows:

$$
\begin{aligned}
ENGS^{(2)}(e^{(2)} \mid e_1^{(1)}, a_2^{(1)}, x_1) \ = \ & \mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_2^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] \\
& -\mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_1^{(2)}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right].
\end{aligned} \qquad (7.23)
$$

Thus, given the inspection outcome at t_1 as $X_1 = x_1$, the optimal inspection action at t_2 is determined as

$$
e^{(2)o} = \arg\max_{e^{(2)}} ENGS^{(2)}(e^{(2)} \mid e_1^{(1)}, a_2^{(1)}, x_1). \qquad (7.24)
$$

The consequences of inspection and replacement options are summarized in the upper half ($e^{(1)} = e_1$ case) of Tables 7.1 and 7.2.

7.3.4 Pre-posterior Analysis: Inspection and Maintenance at t_1

Given the posterior expected cost for each possible X_1, we take an expectation of the posterior cost with respect to X_1. The expected cost arises between t_1 and t_3, with the optimal action derived as

$$\begin{aligned}
& \min_{a^{(1)}} \mathbb{E}_{\boldsymbol{Z}_4}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o})\right] \\
= & \min_{a^{(1)}} \Big(C_P, \mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right] F_{\Delta X_{12}}(\rho_F - x_1) \\
& \quad + C_F(1 - F_{\Delta X_{12}}(\rho_F - x_1))\Big), \qquad (7.25)
\end{aligned}$$

where $\mathbb{E}_{\boldsymbol{Z}_4|X_2<\rho_F}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o}, X_2 < \rho_F)\right]$ is the expected cost with $e^{(2)o}$ and $a^{(2)o}$, given $e_1^{(1)}$. The pre-posterior cost is calculated as

$$C^{prepost} = \mathbb{E}_{X_1}\left[\min_{a^{(1)}} \mathbb{E}_{\boldsymbol{Z}_4}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a_2^{(1)}, a^{(2)o})\right]\right]. \qquad (7.26)$$

By calculating a gap between the two expected optimal costs, we can derive the $ENGS^{(1)}$, which is the net benefit of inspection at t_1, for this single-component problem:

$$ENGS^{(1)} = C^{prior} - C^{prepost}. \qquad (7.27)$$

The consequences of inspection and replacement options are summarized in the upper half ($e^{(1)} = e_1$ case) of Tables 7.3 and 7.4.

7.3.5 Computational Algorithm

Same as the Chapter 6, the PMF method is applied to calculating the expected pre-posterior cost for the two-inspection problem. The algorithm for the two-inspection problem with the PMF method is described as Algorithms 13.

Algorithm 13 Two-inspection single-component problem with PMF method

Discretize the field of X_1
for each $x_{1,j}^{disc}$ in J, **do**
  Calculate $f_{X^{disc}}(x_{1,j}^{disc})$
  Discretize the field of ΔX_{12}
  # For replacement decision at t_2 given $e^{(2)} = e_1$
  for each $\Delta x_{12,i}^{disc}$ in I, **do**
    Calculate $f(\Delta x_{12,i}^{disc})$
    Discretize the field of X_3
    Calculate an expected probability of failure, $\mathbb{E}_{M|x_{1,j}^{disc},\Delta x_{12,i}^{disc}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_{1,j}^{disc} - \Delta x_{12,i}^{disc})\right]$,
      using $f(\Delta x_{23} \mid x_{1,j}, \Delta x_{12,i}^{disc})$
    Calculate an expected cost of taking a_2, $C_F \mathbb{E}_{M|x_{1,j}^{disc},\Delta x_{12,i}^{disc}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_{1,j}^{disc} - \Delta x_{12,i}^{disc})\right]$
    Determine $a^{(2)o}$ by comparing the expected costs of taking a_1 and a_2,
     given $x_{1,j}^{disc}$ and $\Delta x_{12,i}^{disc}$
  end for
  Calculate $\mathbb{E}_{\Delta X_{12}}\left[\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4|x_{1,j}^{disc},\Delta x_{12}}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_1^{(2)}, x_{1,j}^{disc}, \Delta x_{12})\right]\right]$
   $\approx \sum_{i\in I} \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_4|x_{1,j}^{disc},\Delta x_{12,i}^{disc}}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_1^{(2)}, x_{1,j}^{disc}, \Delta x_{12,i}^{disc})\right] f(\Delta x_{12,i}^{disc})$
  # For replacement decision at t_2 given $e^{(2)} = e_2$
  Discretize the field of X_3
  Calculate an expected probability of failure, $\dfrac{\mathbb{E}_{M|x_{1,j}^{disc}}\left[F_{\Delta X_{12}}(\rho_F - x_{1,j}^{disc}) - F_{\Delta X_{13}}(\rho_F - x_{1,j}^{disc})\right]}{\mathbb{E}_{M|x_{1,j}^{disc}}\left[F_{\Delta X_{12}}(\rho_F - x_{1,j}^{disc})\right]}$,
   using $f(\Delta x_{12} \mid x_{1,j})$ and $f(\Delta x_{13} \mid x_{1,j})$
  Calculate an expected cost of taking a_2, $C_F \dfrac{\mathbb{E}_{M|x_{1,j}^{disc}}\left[F_{\Delta X_{12}}(\rho_F - x_{1,j}^{disc}) - F_{\Delta X_{13}}(\rho_F - x_{1,j}^{disc})\right]}{\mathbb{E}_{M|x_{1,j}^{disc}}\left[F_{\Delta X_{12}}(\rho_F - x_{1,j}^{disc})\right]}$
  Determine $a^{(2)o}$ by comparing the expected costs of taking a_1 and a_2, given $x_{1,j}^{disc}$
  # For inspection decision at t_2
  Determine $e^{(2)o}$ by comparing the expected costs of taking e_1 and e_2, given $x_{1,j}^{disc}$
end for
# For replacement decision at t_1
Calculate an expected cost of taking a_2, $C_F \mathbb{E}_{M|x_{1,j}^{disc}}\left[1 - F_{\Delta X_{12}}(\rho_F - x_{1,j}^{disc})\right]$
  $+\mathbb{E}_{M|x_{1,j}^{disc}}\left[F_{\Delta X_{12}}(\rho_F - x_{1,j}^{disc})\right] \mathbb{E}_{\boldsymbol{Z}_4|x_{1,j}^{disc}}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a^{(2)o}, x_{1,j}^{disc})\right]$
Determine $a^{(1)o}$ by comparing the expected costs of taking a_1 and a_2, given $x_{1,j}^{disc}$
For expected pre-posterior cost calculation
Calculate $\mathbb{E}_{X_1}\left[\min_{a^{(1)}} \mathbb{E}_{\boldsymbol{Z}_4|x_1}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e^{(2)o}, a^{(2)o}, x_1)\right]\right]$
  $\approx \sum_{j\in J} \mathbb{E}_{X_1}\left[\min_{a^{(1)}} \mathbb{E}_{\boldsymbol{Z}_4|x_{1,j}^{disc}}\left[C(\boldsymbol{a}_1, \boldsymbol{Z}_4 \mid e_1^{(1)}, e_1^{(2)}, x_{1,j}^{disc})\right]\right] f_{X^{disc}}(x_{1,j}^{disc})$

7.3.6 Numerical Example

We demonstrate the VoI-based approach using an example in which we impose several initial settings: $C_F = 100$, $C_P = 10$, $C_I = 3$, $t_1 = 25$, $t_1 = 27$, $t_3 = 30$, $1/\nu^2 = 9$, and $\rho_F = 3.0$. These settings represent the maintenance problem of a feeder channel in the CANDU 600. Based on observed data from the feeder channels, a parameter uncertainty for μ is set as a random variable, whose prior distribution is given as $Iga\,(\mu; 1102, 97.84)$. With the PMF method, Equation (7.27) is numerically solved.

The *ENGS* is derived as a negative value, -0.154, which means that the optimal inspection is e_2, "do not inspect," at t_1. Consequently, the optimal action at t_1 and inspection at t_2 are determined as $a^{(1)o} = a_2$ and $e^{(2)o} = e_1$, respectively. If the inspection outcome is less than 2.65, the optimal action at t_2 is a_2; otherwise $a^{(2)o} = a_1$.

Influence of Re-inspection at t_2

Although at t_1 operators should not inspect the component, the decision-making after taking $e_1^{(1)}$ is analysed to identify the priority of re-inspection at t_2. The $ENGS^{(2)}$ given $e_1^{(1)}$ becomes positive only if the observation outcome at t_1 is between 2.42 and 2.5, and the maximum $ENGS^{(2)}$ is 2.65, obtained for the case of $x_1 = 2.45$. The $ENGS^{(2)}$ without inspection at t_1 (given $e_2^{(1)}$) is calculated as 1.73, which represents the *ENGS* at t_2 for inspection of an as-yet un-inspected component. The re-inspection priority for an inspected component becomes higher than the first inspection for an un-inspected component only if $2.44 < x_1 < 2.47$, which occurs with a probability of 2.55 %. Therefore, the influence of re-inspection at t_2 on the whole decision-making problem is concluded to be limited.

7.4 Multiple-Component Problem

This section expands the model proposed in Section 7.3 to an N-component system problem.

7.4.1 Additional Assumptions

In addition to the assumptions introduced in Section 7.3, several others are adopted in the problem:

- Inspected components at t_1 will not be re-inspected at t_2;
- An upper limit for the inspection sample size exists;
- The sum of the observed degradation level, S_n, is conditionally independent on each component's degradation level, $X(t)$;

The first assumption is supported by the numerical results of the single-component problem in Section 7.3.6. The numerical analysis indicates that only 8.2 % of previously inspected components have higher priority than un-inspected components, so that the first assumption becomes reasonable in light of the second assumption by which operators can inspect only the components with high priority. The second assumption reflects the reality of operations where full-inspection is not feasible because of limited resources.

For simplification, S_n is derived by using the PDF that is conditionally independent on X instead of calculating $S_n = S_{n-1} + X$. This approximation becomes reasonable when the sample size is large enough. Table 7.5 shows the averages and variances of S_n and $S_{n-1} + x_1$ when $x_1 = 2.5$ is given. These two distributions are close enough; the Kullback-Leibler divergence and the histogram intersection, which are indices showing the difference (similarity) between distributions, are calculated at 0.0528 and 87.1 %, respectively. Note that the Kullback-Leibler

Table 7.5: Statistics for S_n and $S_{n-1} + x_1$

	S_n	$S_{n-1} + x_1$
Average	22.68	22.91
Variance	0.67	0.52

divergence was first introduced by Kullback and Leibler (1951) and has been used to evaluate how much one distribution differs from another distribution. For more detail, see Appendix C.

7.4.2 Random Variables

Let us set the vectors of variables for the whole system, $\boldsymbol{U} = \left(S^{(1)}, S^{(2)}\right)$, and variables for a specific component, $\boldsymbol{W} = (X_1, X_2, X_3)$. The random variables that affect the total expected cost are summarized as a vector, $\boldsymbol{Z}_5 = (M, \boldsymbol{U}, W)$. The PDFs for X_1, X_2, and X_3 with only the prior information are the same as the PDFs in Section 7.3.1. The PDFs for $S^{(1)}$ and $S^{(2)}$, respectively given $n^{(2)}$ and $n^{(2)}$, are

$$f_{S^{(1)}}(s^{(1)} \mid \mu) \quad = \quad ga(s^{(1)}; n^{(1)}t_1/\nu^2, 1/\mu\nu^2), \tag{7.28}$$

$$f_{S^{(2)}}(s^{(2)} \mid \mu, s^{(1)}) \quad = \quad ga(s^{(2)}; n^{(2)}t_2/\nu^2, 1/\mu\nu^2). \tag{7.29}$$

With the prior distribution of μ, the joint distribution of $S^{(1)}$, X_1 and μ is formulated as

$$f_{X_1,S^{(1)},M}(x_1, s^{(1)}, \mu) \quad = \quad f_{X_1}(x_1 \mid \mu) f_{S^{(1)}}(s^{(1)} \mid \mu) f_M(\mu). \tag{7.30}$$

The terms including μ can be summarized as an inverse-gamma distribution, $Iga\left(\mu;\alpha+\frac{(n^{(1)}+1)t_1}{\nu^2},\beta+\frac{s^{(1)}+x_1}{\nu^2}\right)$, so that Equation 7.30 can be modified as

$$f_{X_1,S^{(1)},M}(x_1,s^{(1)},\mu) = \frac{\Gamma(\alpha+\frac{(n^{(1)}+1)t_1}{\nu^2})}{\Gamma(\frac{t_1}{\nu^2})\Gamma(\frac{n^{(1)}t_1}{\nu^2})\Gamma(\alpha)}\frac{\left(\frac{x_1}{\nu^2}\right)^{\frac{t_1}{\nu^2}}\left(\frac{s^{(1)}}{\nu^2}\right)^{\frac{n^{(1)}t_1}{\nu^2}}\beta^{\alpha}}{x_1 s^{(1)}\left(\beta+\frac{x_1+s^{(1)}}{\nu^2}\right)^{\alpha+\frac{(n^{(1)}+1)t_1}{\nu^2}}} \cdot Iga\left(\mu;\alpha+\frac{(n^{(1)}+1)t_1}{\nu^2},\beta+\frac{s^{(1)}+x_1}{\nu^2}\right). \quad (7.31)$$

Thus, the marginalized distribution of $S^{(1)}$ and X_1 with respect to μ is

$$\begin{aligned} f_{X_1,S^{(1)}}(x_1,s^{(1)}) &= \int_0^{\infty} f_{X_1,S^{(1)},M}(x_1,s^{(1)},\mu)d\mu \\ &= \frac{\Gamma(\alpha+\frac{(n^{(1)}+1)t_1}{\nu^2})}{\Gamma(\frac{t_1}{\nu^2})\Gamma(\frac{n^{(1)}t_1}{\nu^2})\Gamma(\alpha)}\frac{\left(\frac{x_1}{\nu^2}\right)^{\frac{t_1}{\nu^2}}\left(\frac{s^{(1)}}{\nu^2}\right)^{\frac{n^{(1)}t_1}{\nu^2}}\beta^{\alpha}}{x_1 s^{(1)}\left(\beta+\frac{x_1+s^{(1)}}{\nu^2}\right)^{\alpha+\frac{(n^{(1)}+1)t_1}{\nu^2}}}. \end{aligned} \quad (7.32)$$

After the observation of $S^{(1)}=s^{(1)}$, the posterior distribution of μ becomes

$$f_M(\mu \mid s^{(1)}) = Iga\left(\mu;\alpha+\frac{n^{(1)}t_1}{\nu^2},\beta+\frac{s^{(1)}}{\nu^2}\right). \quad (7.33)$$

With the posterior distribution, the PDFs of X_2 and $S^{(2)}$ can be derived as

$$\begin{aligned}
f_{X_2}(x_2 \mid s^{(1)}) &= \int_0^\infty f_{X_2}(x_2 \mid \mu) f_M(\mu \mid s^{(1)}) d\mu \\
&= \frac{\Gamma(\alpha + \frac{n^{(1)}t_1 + t_2}{\nu^2})}{\Gamma(\frac{t_2}{\nu^2})\Gamma(\alpha + \frac{n^{(1)}t_1}{\nu^2})} \frac{\left(\frac{x_2}{\nu^2}\right)^{\frac{t_2}{\nu^2}} \left(\beta + \frac{s^{(1)}}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1}{\nu^2}}}{x_2 \left(\beta + \frac{s^{(1)} + x_2}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1 + t_2}{\nu^2}}}, \qquad (7.34) \\
f_{S^{(2)}}(s^{(2)} \mid s^{(1)}) &= \int_0^\infty f_{S^{(2)}}(s^{(2)} \mid \mu) f_M(\mu \mid s^{(1)}) d\mu \\
&= \frac{\Gamma(\alpha + \frac{n^{(1)}t_1 + n^{(2)}t_2}{\nu^2})}{\Gamma(\frac{n^{(2)}t_2}{\nu^2})\Gamma(\alpha + \frac{n^{(1)}t_1}{\nu^2})} \frac{\left(\frac{s^{(2)}}{\nu^2}\right)^{\frac{n^{(2)}t_2}{\nu^2}} \left(\beta + \frac{s^{(1)}}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1}{\nu^2}}}{s^{(2)} \left(\beta + \frac{s^{(1)} + s^{(2)}}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1 + n^{(2)}t_2}{\nu^2}}}, \qquad (7.35)
\end{aligned}$$

and similarly, the joint distribution of X_2 and $S^{(2)}$, given $S^{(1)} = s^{(1)}$, is

$$\begin{aligned}
& f_{X_2, S^{(2)} \mid S^{(1)}}(x_2, s^{(2)} \mid s^{(1)}) \\
= & \int_0^\infty f_{X_2}(x_2 \mid \mu) f_{S^{(2)}}(s^{(2)} \mid \mu) f_M(\mu \mid s^{(1)}) d\mu \\
= & \frac{\Gamma(\alpha + \frac{n^{(1)}t_1 + (n^{(2)}+1)t_2}{\nu^2})}{\Gamma(\frac{t_2}{\nu^2})\Gamma(\frac{n^{(2)}t_2}{\nu^2})\Gamma(\alpha + \frac{n^{(1)}t_1}{\nu^2})} \frac{\left(\frac{x_2}{\nu^2}\right)^{\frac{t_2}{\nu^2}} \left(\frac{s^{(2)}}{\nu^2}\right)^{\frac{n^{(2)}t_2}{\nu^2}} \left(\beta + \frac{s^{(1)}}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1}{\nu^2}}}{x_2 s^{(2)} \left(\beta + \frac{s^{(1)} + x_2 + s^{(2)}}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1 + (n^{(2)}+1)t_2}{\nu^2}}}. \qquad (7.36)
\end{aligned}$$

Once operators observe $s^{(1)}$ and $s^{(2)}$, the PDF of μ is further updated as

$$f_M(\mu \mid s^{(1)}, s^{(2)}) = Iga\left(\mu; \alpha + \frac{n^{(1)}t_1 + n^{(2)}t_2}{\nu^2}, \beta + \frac{s^{(1)} + s^{(2)}}{\nu^2}\right). \qquad (7.37)$$

The PDF of X_3, which will be used for calculating the probability of failure between t_2 and t_3, is derived as follows:

$$
\begin{aligned}
f_{X_3|S^{(1)},S^{(2)}}(x_3 \mid s^{(1)}, s^{(2)}) &= \int_0^{\infty} f_{X_3}(x_3 \mid \mu) f_M(\mu \mid s^{(1)}, s^{(2)}) d\mu \\
&= \frac{\Gamma(\alpha + \frac{n^{(1)}t_1 + n^{(2)}t_2 + t_3}{\nu^2})}{\Gamma(\frac{t_3}{\nu^2})\Gamma(\alpha + \frac{n^{(1)}t_1 + n^{(2)}t_2}{\nu^2})} \frac{\left(\frac{x_3}{\nu^2}\right)^{\frac{t_3}{\nu^2}} \left(\beta + \frac{s^{(1)}+s^{(2)}}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1 + n^{(2)}t_2}{\nu^2}}}{x_3 \left(\beta + \frac{s^{(1)}+s^{(2)}}{\nu^2}\right)^{\alpha + \frac{n^{(1)}t_1 + n^{(2)}t_2 + t_3}{\nu^2}}}
\end{aligned}
\tag{7.38}
$$

7.4.3 Baseline Case: No Inspection at t_1

The total cost depends on inspection and replacement actions, $\boldsymbol{a}_2 = (e^{(1)}, e^{(2)}, n^{(1)}, n^{(2)}, a^{(1)}, a^{(2)})$, and random variables, $\boldsymbol{Z}_5 = (M, \boldsymbol{U}, W)$. As in the single-component problem, the optimization analysis without the first sampling inspection is actually the same as the pre-posterior analysis in Section 6.3, except for the shift of the time horizon and the additional maintenance decision at t_1. In this section, the expected cost without any inspection is first introduced, and the case with sampling inspection of $n^{(2)}$ components is analysed for both un-inspected and inspected components.

No Inspection at Either t_1 or t_2

If operators have no observation outcome for either t_1 or t_2, a parameter uncertainty cannot be updated. Thus, the prior distribution is used for calculating expected cost. The expected optimal cost with no inspection at either t_1 or t_2 is calculated:

$$
\begin{aligned}
&\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|X_2<\rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, n^{(1)} = 0, n^{(2)} = 0, a_2^{(1)}) \right] \\
= \; &\min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_M \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]} \right).
\end{aligned}
\tag{7.39}
$$

Un-inspected Components at t_2

Suppose operators inspect $n^{(2)}$ components at t_2 and obtain a sum of the inspected degradation level at $s^{(2)}$. The optimal expected cost of a component with no inspection at either t_1 or t_2 is calculated thus:

$$\begin{aligned} & \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|S^{(2)},X_2<\rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, n^{(1)} = 0, n^{(2)}, a_2^{(1)}, s^{(2)}) \right] \\ = \; & \min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_{M|S^{(2)}} \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]} \right). \end{aligned} \tag{7.40}$$

The expected pre-posterior cost becomes

$$\begin{aligned} & \mathbb{E}_{S^{(2)}} \left[\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|S^{(2)},X_2<\rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, n^{(1)} = 0, n^{(2)}, a_2^{(1)}, s^{(2)}) \right] \right] \\ = \; & \mathbb{E}_{S^{(2)}} \left[\min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_{M|S^{(2)}} \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]} \right) \right]. \end{aligned} \tag{7.41}$$

Inspected Components at t_2

The expected cost of a component with inspection at t_2, given $X_2 = x_2$ and $S^{(2)} = s^{(2)}$, is

$$\begin{aligned} & \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|X_2,S^{(2)}} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_1^{(2)}, n^{(1)} = 0, n^{(2)}, a_2^{(1)}, x_2, s^{(2)}) \right] \\ = \; & C_I + \min_{a^{(2)}} \left(C_P, C_F \mathbb{E}_{M|S^{(2)}} \left[1 - F_{\Delta X_{23}}(\rho_F - x_2) \right] \right) \end{aligned} \tag{7.42}$$

The expected cost unconditional on X_2 and $S^{(2)}$ becomes

$$\begin{aligned} & \mathbb{E}_{X_2,S^{(2)}|X_2<\rho_F} \left[\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|X_2,S^{(2)}} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_1^{(2)}, n^{(1)} = 0, n^{(2)}, a_2^{(1)}, x_2) \right] \right] \\ = \; & C_I + \mathbb{E}_{S^{(2)},X_2|X_2<\rho_F} \left[\min_{a^{(2)}} \left(C_P, C_F \mathbb{E}_{M|S^{(2)}} \left[1 - F_{\Delta X_{23}}(\rho_F - x_2) \right] \right) \right], \end{aligned} \tag{7.43}$$

Table 7.6: Costs arising between t_2 and t_3 given $n^{(1)} = 0$, $a^{(1)} = a_2$, and $X_2 < \rho_F$: Prior analysis

$n^{(2)}$	$s^{(2)}$	$e^{(2)}$	x_2	$a^{(2)}$	State	Cost	Probability
$n^{(2)}$	✓	e_1	✓	a_1		$C_I + C_P$	1
				a_2	Fail	$C_I + C_F$	$\mathbb{E}_{M\mid S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
					Survive	C_I	$\mathbb{E}_{M\mid S^{(2)}}\left[F_{\Delta X_{23}}(\rho_F - x_2)\right]$
		e_2	×	a_1		C_P	1
				a_2	Fail	C_F	$\frac{\mathbb{E}_{M\mid S^{(2)}}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}$
					Survive	0	$\frac{\mathbb{E}_{M\mid S^{(2)}}\left[F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}$
0	×	e_2	×	a_1		C_P	1
				a_2	Fail	C_F	$\frac{\mathbb{E}_M\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}$
					Survive	0	$\frac{\mathbb{E}_M\left[F_{X_3}(\rho_F)\right]}{\mathbb{E}_M\left[F_{X_2}(\rho_F)\right]}$

Note that this expected optimal cost is for a case in which the component will not fail before t_2, a case whose probability is $F_{X_2}(\rho_F)$. The consequences of inspection and replacement options are summarized in Tables 7.6 and 7.7.

Optimal Inspection at t_2

Based on the derived expected costs, the expected cost for each inspection option $e^{(2)}$ is derived as in Table 7.8. Thus, the $ENGS^{(2)}$ is derived as follows:

$$\begin{aligned}
& ENGS^{(2)}(n^{(2)} \mid n^{(1)} = 0, a_2^{(1)}) \\
= \; & N \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5 \mid X_2 < \rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, n^{(1)} = 0, n^{(2)} = 0, a_2^{(1)}) \right] \\
& -(N - n^{(2)}) \mathbb{E}_{S^{(2)}} \left[\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5 \mid S^{(2)}, X_2 < \rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, n^{(1)} = 0, n^{(2)}, a_2^{(1)}, s^{(2)}) \right] \right] \\
& -n^{(2)} \mathbb{E}_{X_2, S^{(2)} \mid X_2 < \rho_F} \left[\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5 \mid X_2, S^{(2)}} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_1^{(2)}, n^{(1)} = 0, n^{(2)}, a_2^{(1)}, x_2) \right] \right]
\end{aligned} \tag{7.44}$$

Table 7.7: Costs arising between t_1 and t_3 given $n^{(1)} = 0$, $a^{(1)} = a_2$, and $X_2 < \rho_F$: Prior analysis

$n^{(2)}$	$s^{(2)}$	$e^{(2)}$	x_2	$a^{(2)}$	$\mathbb{E}_{\boldsymbol{Z}_5 \mid X_2 < \rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid a_2^{(1)}, X_2 < \rho_F) \right]$
$n^{(2)}$	✓	e_1	✓	a_1	$C_I + C_P$
				a_2	$C_I + C_F \mathbb{E}_{M \mid S^{(2)}} \left[1 - F_{\Delta X_{23}}(\rho_F - x_2) \right]$
		e_2	×	a_1	C_P
				a_2	$C_F \frac{\mathbb{E}_{M \mid S^{(2)}} \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]}$
0	×	e_2	×	a_1	C_P
				a_2	$C_F \frac{\mathbb{E}_M \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]}$

Table 7.8: Total costs for inspected and un-inspected components: Prior analysis

$n^{(2)}$	$s^{(2)}$	$e^{(2)}$	x_2	$\mathbb{E}_{\boldsymbol{Z}_5 \mid X_2 < \rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid a_2^{(1)}, a^{(2)o}, X_2 < \rho_F) \right]$
$n^{(2)}$	✓	e_1	✓	$n^{(2)} \left\{ C_I + \mathbb{E}_{S^{(2)}, X_2 \mid X_2 < \rho_F} \left[\min_{a^{(2)}} \left(C_P, C_F \mathbb{E}_{M \mid S^{(2)}} \left[1 - F_{\Delta X_{23}}(\rho_F - x_2) \right] \right) \right] \right\}$
		e_2	×	$(N - n^{(2)}) \mathbb{E}_{S^{(2)}} \left[\min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_{M \mid S^{(2)}} \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]} \right) \right]$
0	×	e_2	×	$N \min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_M \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]} \right)$

The optimal inspection action at t_2 is determined as

$$n^{(2)o} = \arg\max_{n^{(2)}} ENGS^{(2)}(n^{(2)} \mid n^{(1)} = 0, a_2^{(1)}). \tag{7.45}$$

The average cost with the optimal inspection at t_2 is

$$\begin{aligned} & \mathbb{E}_{\boldsymbol{Z}_5} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e^{(2)o}, n^{(1)} = 0, n^{(2)o}, a_2^{(1)}, a^{(2)o}) \right] \\ = \; & \frac{1}{N} \left\{ (N - n^{(2)}) \mathbb{E}_{S^{(2)}} \left[\min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_{M \mid S^{(2)}} \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_M \left[F_{X_2}(\rho_F) \right]} \right) \right] \right. \\ & \left. + n^{(2)} \left\{ C_I + \mathbb{E}_{S^{(2)}, X_2 \mid X_2 < \rho_F} \left[\min_{a^{(2)}} \left(C_P, C_F \mathbb{E}_{M \mid S^{(2)}} \left[1 - F_{\Delta X_{23}}(\rho_F - x_2) \right] \right) \right] \right\} \right\}. \end{aligned} \tag{7.46}$$

At the maintenance stage at t_1, the operators choose an action between a_1, replacement, and a_2, no-action, with which the expected cost is lower than the cost with the other option. The "prior" cost of the two-inspection problem is

$$\begin{aligned} C^{prior} &= \min_{a^{(1)}} \Big(C_P, F_{X_2}(\rho_F) \mathbb{E}_{\boldsymbol{Z}_5} \Big[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e^{(2)o}, n^{(1)} = 0, n^{(2)o}, a_2^{(1)}, a^{(2)o}) \Big] \\ &\qquad + C_F(1 - F_{X_2}(\rho_F)) \Big) . \end{aligned} \tag{7.47}$$

7.4.4 Posterior Analysis: Inspection and Maintenance at t_2

Suppose that operators have obtained an outcome for a sampling inspection, $s^{(1)}$ and x_1. The operators will then choose the best action according to a comparison between the expected costs, using updated information about M, X_2, X_3, and $S^{(2)}$. The consequences of all inspection and action options are summarized in Tables 7.9, 7.10, and 7.11.

Un-inspected Components at Either t_1 or t_2

Even when operators are going to inspect $n^{(1)}$ and $n^{(2)}$ components at t_1 and t_2, respectively, they still have $N - n^{(1)} - n^{(2)}$ un-inspected components at either t_1 or t_2 if $N > n^{(1)} + n^{(2)}$. The optimal expected cost given $a^{(1)} = a_2$, $e^{(1)} = e_2$, $e^{(2)} = e_2$, and $x_2 < \rho_F$, is calculated as

$$\begin{aligned} &\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5 | S^{(1)}, S^{(2)}, X_2 < \rho_F} \Big[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, a_2^{(1)}, n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)}) \Big] \\ = &\min_{a^{(2)}} \left(C_P, C_F \frac{\mathbb{E}_{M|S^{(1)}, S^{(2)}} \left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F) \right]}{\mathbb{E}_{M|S^{(1)}} \left[F_{X_2}(\rho_F) \right]} \right). \end{aligned} \tag{7.48}$$

Inspected Components at t_2 Only

For $n^{(2)}$ components, given $a^{(1)} = a_2$, $e^{(1)} = e_2$, $e^{(2)} = e_1$, and $X_2 = x_2 < \rho_F$, the optimal expected cost is

$$\begin{aligned} &\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|S^{(1)},S^{(2)},X_2}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_1^{(2)}, a_2^{(1)},, n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)}, x_2)\right] \\ = &\left(C_I + \mathbb{E}_{S^{(2)},X_2|S^{(1)},X_2<\rho_F}\left[\min_{a^{(2)}}(C_P, C_F \mathbb{E}_{M|S^{(1)},S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]\right]\right). \end{aligned} \tag{7.49}$$

From Equations (7.48) and (7.49), the average of the optimal expected cost, given $e^{(1)} = e_2$, is calculated as

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{Z}_5|S^{(1)},S^{(2)},X_2<\rho_F}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, a_2^{(1)}, a^{(2)o},, n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)})\right] \\ = &\frac{1}{N - n^{(1)}}\Big\{(N - n^{(1)} - n^{(2)}) \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|S^{(1)},S^{(2)},X_2<\rho_F}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, a_2^{(1)}, n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)})\right] \\ &+ n^{(2)} \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|S^{(1)},S^{(2)},X_2}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_1^{(2)}, a_2^{(1)},, n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)}, x_2)\right]\Big\}. \end{aligned} \tag{7.50}$$

Inspected Components at t_1 Only

For $n^{(1)}$ components, when $a^{(1)} = a_2$, $e^{(1)} = e_1$, $e^{(2)} = e_2$, and $X_2 = x_2 < \rho_F$ are given, the optimal expected cost is

$$\begin{aligned} &\min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|X_2}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_1^{(1)}, e_2^{(2)}, a_2^{(1)},, n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)}, x_1)\right] \\ = &\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M|S^{(1)},S^{(2)}}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M|S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}\right) \end{aligned} \tag{7.51}$$

Inspection Decision at t_2

In the two-inspection problem, the second inspection decision is included in the posterior analysis part. Based on the derived expected costs, the $ENGS^{(2)}$ is derived as follows:

$$\begin{aligned}
& ENGS^{(2)}(n^{(2)} \mid n^{(1)}, a_2^{(1)}) \\
= \; & N \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|S^{(1)},X_1,X_2<\rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, e_2^{(2)}, a_2^{(1)},) \right] \\
& -(N - n^{(1)}) \mathbb{E}_{\boldsymbol{Z}_5|S^{(1)},S^{(2)},X_2<\rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, a_2^{(1)}, a^{(2)o}, , n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)}) \right] \\
& -n^{(1)} \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|X_2} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_1^{(1)}, e_2^{(2)}, a_2^{(1)}, , n^{(1)}, n^{(2)}, s^{(1)}, s^{(2)}, x_1) \right]. \qquad (7.52)
\end{aligned}$$

The optimal sample size at t_2 is determined to be

$$n^{(2)o} = \arg\max_{n^{(2)}} ENGS^{(2)}(n^{(2)} \mid n^{(1)}, a_2^{(1)}, a^{(2)o}). \qquad (7.53)$$

Table 7.9: Costs arising between t_2 and t_3 given $n^{(1)}$, $s^{(1)}$, $a^{(1)} = a_2$ and $X_2 < \rho_F$: Posterior analysis

$n^{(1)}$	$s^{(1)}$	$e^{(1)}$	x_1	$a^{(1)}$	State	$n^{(2)}$	$s^{(2)}$	$e^{(2)}$	x_2	$a^{(2)}$	State	Cost	Probability
$n^{(1)}$	✓	e_1	✓	a_1								—	—
				a_2	Fail							—	—
					Survive	$n^{(2)}$	✓	e_1	✓	a_1		$C_I + C_P$	1
										a_2	Fail	$C_I + C_F$	$\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
											Survive	C_I	$\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[F_{\Delta X_{23}}(\rho_F - x_2)\right]$
								e_2	×	a_1		C_P	1
										a_2	Fail	C_F	$\frac{\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[F_{\Delta X_{12}}(\rho_F-x_1)-F_{\Delta X_{13}}(\rho_F-x_1)\right]}{\mathbb{E}_{M\|S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F-x_1)\right]}$
											Survive	0	$\frac{\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[F_{\Delta X_{13}}(\rho_F-x_1)\right]}{\mathbb{E}_{M\|S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F-x_1)\right]}$
						0	×	e_2	×	a_1		C_P	1
										a_2	Fail	C_F	$\frac{\mathbb{E}_{M\|S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F-x_1)-F_{\Delta X_{13}}(\rho_F-x_1)\right]}{\mathbb{E}_{M\|S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F-x_1)\right]}$
											Survive	0	$\frac{\mathbb{E}_{M\|S^{(1)}}\left[F_{\Delta X_{13}}(\rho_F-x_1)\right]}{\mathbb{E}_{M\|S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F-x_1)\right]}$
		e_2	×	a_1								—	—
				a_2	Fail							—	—
					Survive	$n^{(2)}$	✓	e_1	✓	a_1		$C_I + C_P$	1
										a_2	Fail	$C_I + C_F$	$\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
											Survive	C_I	$\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[F_{\Delta X_{23}}(\rho_F - x_2)\right]$
								e_2	×	a_1		C_P	1
										a_2	Fail	C_F	$\frac{\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[F_{X_2}(\rho_F)-F_{X_3}(\rho_F)\right]}{F_{X_2}(\rho_F)}$
											Survive	0	$\frac{\mathbb{E}_{M\|S^{(1)},S^{(2)}}\left[F_{X_3}(\rho_F)\right]}{F_{X_2}(\rho_F)}$
						0	×	e_2	×	a_1		C_P	1
										a_2	Fail	C_F	$\frac{\mathbb{E}_{M\|S^{(1)})}\left[F_{X_2}(\rho_F)-F_{X_3}(\rho_F)\right]}{F_{X_2}(\rho_F)}$
											Survive	0	$\frac{\mathbb{E}_{M\|S^{(1)}}\left[F_{X_3}(\rho_F)\right]}{F_{X_2}(\rho_F)}$

Table 7.10: Costs arising between t_2 and t_3 given $n^{(1)}$, $s^{(1)}$, $a^{(1)} = a_2$ and $X_2 < \rho_F$: Posterior analysis

$n^{(1)}$	$s^{(1)}$	$e^{(1)}$	x_1	$a^{(1)}$	State	$n^{(2)}$	$s^{(2)}$	$e^{(2)}$	x_2	$a^{(2)}$	$\mathbb{E}_{\boldsymbol{Z}_5 \mid X_2 < \rho_F}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid a_2^{(1)}, X_2 < \rho_F)\right]$
$n^{(1)}$	✓	e_1	✓	a_1							—
				a_2	Fail						—
					Survive	$n^{(2)}$	✓	e_2	×	a_1	C_P
										a_2	$C_F \frac{\mathbb{E}_{M\mid S^{(1)},S^{(2)}}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}$
						0	×	e_2	×	a_1	C_P
										a_2	$C_F \frac{\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}$
		e_2	×	a_1							—
				a_2	Fail						—
					Survive	$n^{(2)}$	✓	e_1	✓	a_1	$C_I + C_P$
										a_2	$C_I + C_F \mathbb{E}_{M\mid S^{(1)},S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]$
								e_2	×	a_1	C_P
										a_2	$C_F \frac{\mathbb{E}_{M\mid S^{(1)},S^{(2)}}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{F_{X_2}(\rho_F)}$
						0	×	e_2	×	a_1	C_P
										a_2	$C_F \frac{\mathbb{E}_{M\mid S^{(1)})}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{F_{X_2}(\rho_F)}$

Table 7.11: Total costs for inspected and un-inspected components: Posterior analysis

$n^{(1)}$	$s^{(1)}$	$e^{(1)}$	x_1	$a^{(1)}$	State	$n^{(2)}$	$s^{(2)}$	$e^{(2)}$	x_2	Total cost
$n^{(1)}$	✓	e_1	✓	a_1						—
				a_2	Fail					—
					Survive	$n^{(2)}$	✓	e_2	×	$n^{(1)} \mathbb{E}_{S^{(2)}\mid S^{(1)}}\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(1)},S^{(2)}}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}\right)\right]$
						0	×	e_2	×	$n^{(1)}\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}\right)\right]$
		e_2	×	a_1						—
				a_2	Fail					—
					Survive	$n^{(2)}$	✓	e_1	✓	$n^{(2)}\Big(C_I + \mathbb{E}_{S^{(2)},X_2\mid S^{(1)},X_2<\rho_F}\big[\min_{a^{(2)}}(C_P,$ $C_F \mathbb{E}_{M\mid S^{(1)},S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]\big]\Big)$
								e_2	×	$(N - n^{(1)} - n^{(2)})$ $\mathbb{E}_{S^{(2)}\mid S^{(1)}}\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(1)},S^{(2)}}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{X_2}(\rho_F)\right]}\right)\right]$
						0	×	e_2	×	$(N - n^{(1)})\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(1)})}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{X_2}(\rho_F)\right]}\right)\right]$

7.4.5 Pre-posterior Analysis

In the pre-posterior analysis, the expected costs with respect to the two random variables, X_1 and $S^{(1)}$, are considered for both un-inspected and inspected components. The detailed consequence of each option is shown in Tables 7.12 and 7.13.

The pre-posterior cost for an un-inspected component is calculated as

$$\begin{aligned} C_{un}^{prepost} &= \min_{a^{(1)}} \Big(C_P, C_F \mathbb{E}_{M|S^{(1)}} \left[1 - F_{X_2}(\rho_F)\right] \\ &\quad + \mathbb{E}_{M|S^{(1)}} \left[F_{X_2}(\rho_F)\right] \mathbb{E}_{\boldsymbol{Z}_5|X_2<\rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, n^{(2)o}, a^{(2)o})) \right] \Big) \end{aligned} \tag{7.54}$$

The pre-posterior cost for an inspected component is derived as

$$C_{insp}^{prepost} = \min_{a^{(2)}} \mathbb{E}_{\boldsymbol{Z}_5|X_2<\rho_F} \left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid e_2^{(1)}, n^{(2)o}, a^{(2)o}, x_1) \right]. \tag{7.55}$$

By calculating a gap between the two expected optimal costs, we can derive the *ENGS* for this single-component problem:

$$ENGS^{(1)}(n^{(1)}) = NC_{prior} - (N - n^{(1)}) C_{un}^{prepost} - n^{(1)} C_{insp}^{prepost}. \tag{7.56}$$

Finally, the optimal sample size at t_1 is determined to be

$$n^{(1)o} = \arg\max_{n^{(1)}} ENGS^{(1)}(n^{(1)}). \tag{7.57}$$

Table 7.12: Costs arising between t_1 and t_3: Pre-posterior analysis

$n^{(1)}$	$s^{(1)}$	$e^{(1)}$	x_1	$a^{(1)}$	$\mathbb{E}_{\boldsymbol{Z}_5}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid n^{(1)}, n^{(2)o}, a^{(2)o}\right]$
$n^{(1)}$	✓	e_1	✓	a_1	$C_I + C_P$
				a_2	$C_I + C_F \mathbb{E}_{M\mid S^{(1)}}\left[1 - F_{\Delta X_{12}}(\rho_F - x_1)\right]$ $+\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right] \mathbb{E}_{S^{(2)}_{n^{(2)o}}\mid S^{(1)}}\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(1)}, S^{(2)}_{n^{(2)o}}}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1)\right]}\right)\right]$
		e_2	×	a_1	C_P
				a_2	$C_F \mathbb{E}_{M\mid S^{(1)}}\left[1 - F_{X_2}(\rho_F)\right]$ $+\frac{1}{N-n^{(1)}}\left\{n^{(2)o}\left(C_I + \mathbb{E}_{S^{(2)}, X_2\mid S^{(1)}, X_2<\rho_F}\left[\min_{a^{(2)}}(C_P, C_F \mathbb{E}_{M\mid S^{(1)}, S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]\right]\right)\right\}$ $+(N - n^{(1)} - n^{(2)o})\mathbb{E}_{S^{(2)}\mid S^{(1)}}\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(1)}, S^{(2)}}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{X_2}(\rho_F)\right]}\right)\right]\Big\}$

Table 7.13: Costs arising between t_1 and t_3 given optimal action $a^{(1)o}$: Pre-posterior analysis

$n^{(1)}$	$s^{(1)}$	$e^{(1)}$	x_1	$\mathbb{E}_{\boldsymbol{Z}_5}\left[C(\boldsymbol{a}_2, \boldsymbol{Z}_5 \mid n^{(2)o}, a^{(1)}, a^{(2)o}\right]$
$n^{(1)}$	✓	e_1	✓	$C_I + \min_{a^{(1)}}\Big(C_P, C_F \mathbb{E}_{M\mid S^{(1)}}\left[1 - F_{\Delta X_{12}}(\rho_F - x_1)\right]$ $+\mathbb{E}_{M\mid S^{(1)}}\left[F_{\Delta X_{12}}(\rho_F - x_1) \min_{a^{(2)}}\left(C_P, C_F \mathbb{E}_{M\mid S^{(1)}, S^{(2)}}\left[F_{\Delta X_{12}}(\rho_F - x_1) - F_{\Delta X_{13}}(\rho_F - x_1)\right]\right]\right)\Big)$
		e_2	×	$\min_{a^{(1)}}\Big(C_P, C_F \mathbb{E}_{M\mid S^{(1)}}\left[1 - F_{X_2}(\rho_F)\right]$ $+\frac{1}{N-n^{(1)}}\left\{n^{(2)o}\left(C_I + \mathbb{E}_{S^{(2)}, X_2\mid S^{(1)}, X_2<\rho_F}\left[\min_{a^{(2)}}(C_P, C_F \mathbb{E}_{M\mid S^{(1)}, S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]\right]\right)\right\}$ $+(N - n^{(1)} - n^{(2)o})\mathbb{E}_{S^{(2)}\mid S^{(1)}}\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(1)}, S^{(2)}}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_{M\mid S^{(1)}}\left[F_{X_2}(\rho_F)\right]}\right)\right]\Big\}\Big)$
0	×	e_2	×	$\min_{a^{(1)}}\Big(C_P, C_F \mathbb{E}_{M\mid S^{(1)}}\left[1 - F_{X_2}(\rho_F)\right]$ $+\frac{1}{N}\left\{n^{(2)o}\left\{C_I + \mathbb{E}_{S^{(2)}, X_2\mid X_2<\rho_F}\left[\min_{a^{(2)}}\left(C_P, C_F \mathbb{E}_{M\mid S^{(2)}}\left[1 - F_{\Delta X_{23}}(\rho_F - x_2)\right]\right)\right]\right\}\right.$ $+(N - n^{(2)o})\mathbb{E}_{S^{(2)}}\left[\min_{a^{(2)}}\left(C_P, C_F \frac{\mathbb{E}_{M\mid S^{(2)}}\left[F_{X_2}(\rho_F) - F_{X_3}(\rho_F)\right]}{\mathbb{E}_{M}\left[F_{X_2}(\rho_F)\right]}\right)\right]\Big\}$

7.5 Numerical Example

We next demonstrate the VoI-based approach using an example in which we impose several initial settings: $N = 319$, $C_F = 100$, $C_P = 10$, $C_I = 1$, $t_1 = 25$, $t_2 = 27$, $t_3 = 30$, $1/\nu^2 = 9$, and $\rho_F = 3.0$. Note that the population size represents the number of un-inspected components of a 380-component system. Let us assume that the prior distribution for μ is given as $Iga\,(\mu; 1102, 97.84)$, derived from observed data. Let us set the options for inspection sample size at 10, 20, 30, 40, and 50; the sample size of each inspection is limited to 50. The simulation algorithms adopt the PMF method introduced in Chapter 6.

7.5.1 Illustration of Proposed Policy

Derived from the backward induction illustrated in Section 7.4, an optimal policy is proposed. The best sample size at t_1 is $n^{(1)o} = 40$ with its *ENGS* of 67.77. The optimal action, $a^{(1)o}$, after inspection of 40 components, depends on the inspection's outcome. Figure 7.2 illustrates the optimal action for inspected or un-inspected components and its ranges of X_1 and $S^{(1)}$. As for the 279 un-inspected components, operators will replace them if the sum of the observed degradation level is more than 96.0. For the 40 inspected components, operators will decide on an action based on both X_1 and $S^{(1)}$. At t_2, the optimal inspection is determined relying on $S^{(1)}$. If $90.5 < S^{(1)} \leq 96.0$, the optimal sample size at t_2 is 50; otherwise, no inspection is required. The optimal action at t_2, $a^{(2)}$, is derived based on random variables $\boldsymbol{Z}_5$, and operators determine policy-$\boldsymbol{a}$.

Table 7.14: *ENGS*$^{(1)}$ of the two-inspection problem for each sample size

Sample size	*ENGS*
10	51.98
20	58.25
30	59.67
40	59.99
50	59.65

Table 7.15: *ENGS* of one-inspection problem for each sample size at t_1

Inspection	$n^{(1)o}$	$n^{(2)o}$	Total expected cost	Benefit of inspection
t_1 and t_2	40	0 or 50	1029.9	112.6
t_1	20	0	1088.1	54.4
t_2	0	30	1089.8	52.7
No inspection	0	0	1142.5	—

7.5.2 Comparison with One-Inspection Problem

The results of the two-inspection problem are compared with those for a one- or no-inspection problem, as summarized in Table 7.15. Note that the total expected costs are compared instead of the *ENGS*, because the baseline of the analysis (prior analysis) of each problem is different. The results suggest several insights:

- Inspection at t_1 is more effective than one at t_2 if only one inspection is allowed;
- The benefit of inspection at both t_1 and t_2 is more than the sum of the benefits of separate inspections at t_1 and t_2;
- Synergy between inspection at t_1 and t_2 is indicated.

Note that the third insight is obvious because the second insight is observed, although the marginal influence of reducing parameter uncertainty decreases as the total sample size, $n^{(1)}+n^{(2)}$, increases.

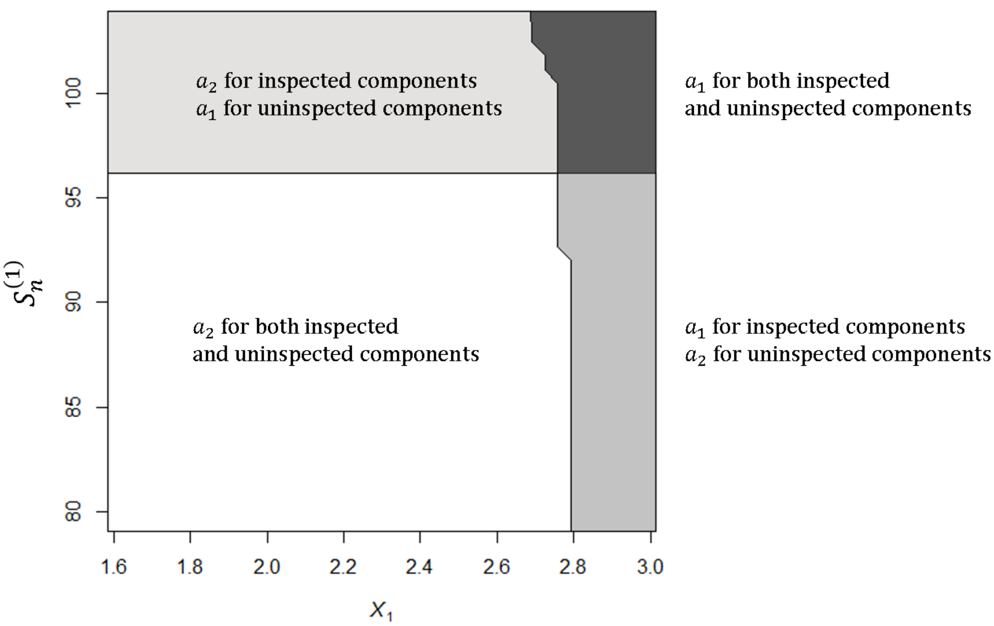

Figure 7.2: Optimal action, $a^{(1)o}$, for each combination of observation outcomes of X_1 and $S^{(1)}$ given $n^{(1)o} = 40$

7.5.3 Sensitivity Analysis

With the same initial settings except for a focusing parameter, the optimal sample sizes are derived plus the sensitivities of C_I, C_P, C_F, and the parameters of prior distribution, (α, β), are discussed. To illustrate the influence of the second inspection at t_2, the sensitivity analysis for the two-inspection problem is compared with the analysis for the one-inspection (t_1) problem. Note that the results of a two-inspection problem always show lower expected costs than the results of a one-inspection problem, because a one-inspection problem is just a special case of a two-inspection problem.

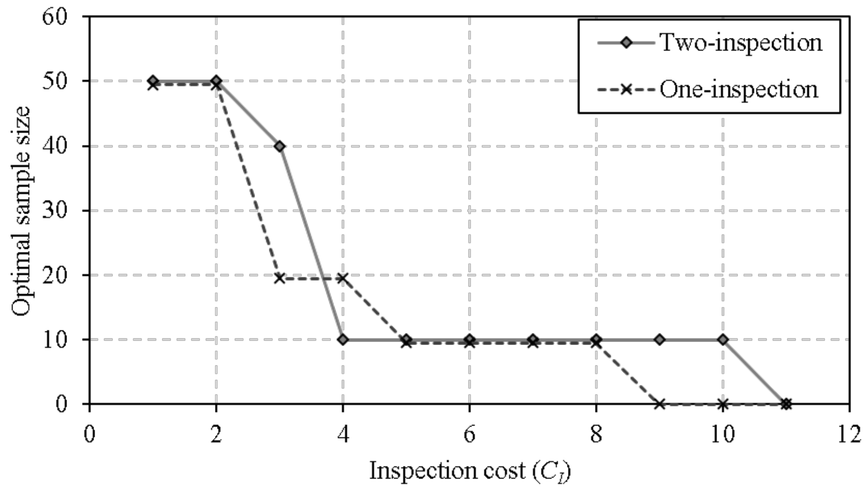

Figure 7.3: Optimal sample size with different C_I ($C_P = 10$ and $C_F = 100$)

Impact of Cost Balance

Figure 7.3 illustrates the *ENGS*s with each C_I. Although the general trend is similar to the results of the one-inspection problem shown in Figure 6.7, the detailed behaviour of the *ENGS* provides several insights into the difference between the two problems. In the cases where operators have a chance to inspect at t_2, a moderate inspection option ($10 \leq n \leq 40$) becomes optimal when $3 \geq C_I \geq 10$, which is slightly wider than the range calculated with the one-inspection at t_1 case, $3 \geq C_I \geq 8$.

The influence of the replacement cost, C_P, and the failure cost, C_F, are summarized in Figures 7.4 and 7.5, respectively. For the sensitivity analysis on C_P, the optimal sample size peaks at around $C_P = 10$ and decreases as C_P increases. The moderate inspection options ($n^{(1)} = 10, 20, 30$ or 40) become optimal when C_P is between 5 and 35. In contrast to the results for C_P, the optimal sample size monotonically and non-linearly increases from 0 to 50 as the C_F changes from 50 to 110. In the two-inspection case, the range of C_P or C_F, where a moderate

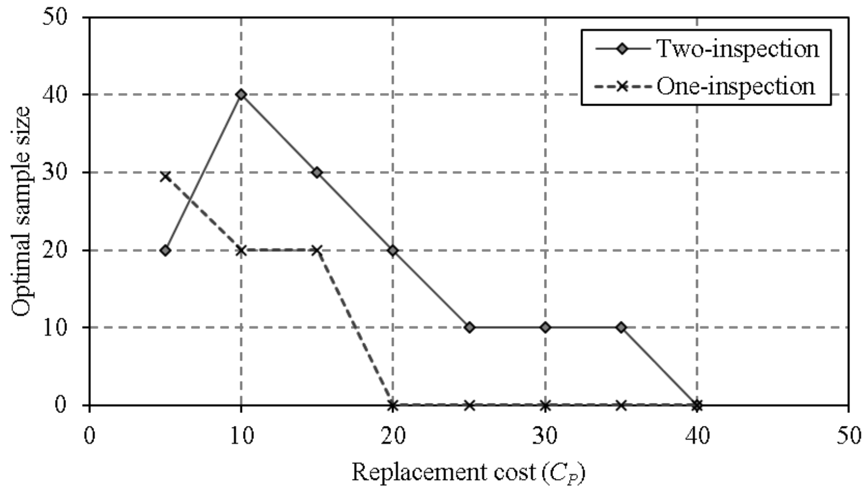

Figure 7.4: Optimal sample size with different C_P ($C_I = 3$ and $C_F = 100$)

inspection option ($10 \leq n \leq 40$) becomes optimal, is wider than their range in the one-inspection at t_1 case.

Impact of Prior Information

The influence of the amount of prior information, which represents the precision of prior distribution, is analysed as in Figure 7.6. With the fixed ratio of α/β, α and β are changed from 10 % to 250 % of the original values, $(\alpha, \beta) = (1102, 97.8)$. Similar to the results of the one-inspection problem, the less prior information operators have, the more the sample size is optimal, although the optimal sample size can reach only 50, as that is the upper limit possible. The two-inspection case has positive *ENGS* even if the operators have at most 2.4 times of the original prior information, whereas the one-inspection case has positive *ENGS* when the prior information is equal to or less than 1.6 times the original.

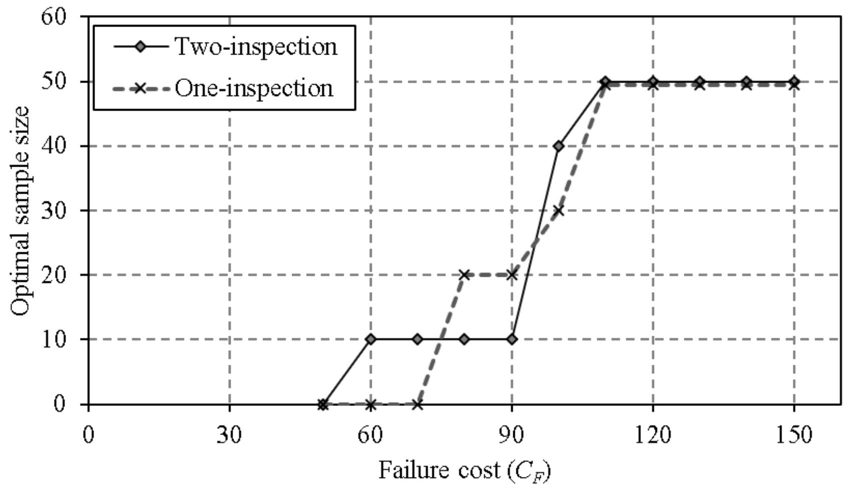

Figure 7.5: Optimal sample size with different C_F ($C_I = 3$ and $C_P = 10$)

7.5.4 Discussion

Major insights from the numerical-analysis results are as follows:

- The chance of a second inspection enhances the value of inspection at t_1;
- Synergy between inspections at t_1 and t_2 is indicated;
- Regardless of the large population size, moderate inspection options, $n^{(1)} = 10, 20, 30$, or 40, tend to be optimal within a wider range of C_I, C_P, and C_F.

These insights indicate several findings for a general multiple-inspection problem when they are combined with the insights in Section 6.5.2. When the inspection cost is low enough to obtain benefit from reducing only aleatory uncertainty, at each inspection outage, operators should inspect samples of maximum size until all components have been inspected. Once all un-inspected components are inspected, the operators need to calculate the *ENGS* for each component, ignoring parameter uncertainties, and inspect only the components that have positive *ENGS*. When the

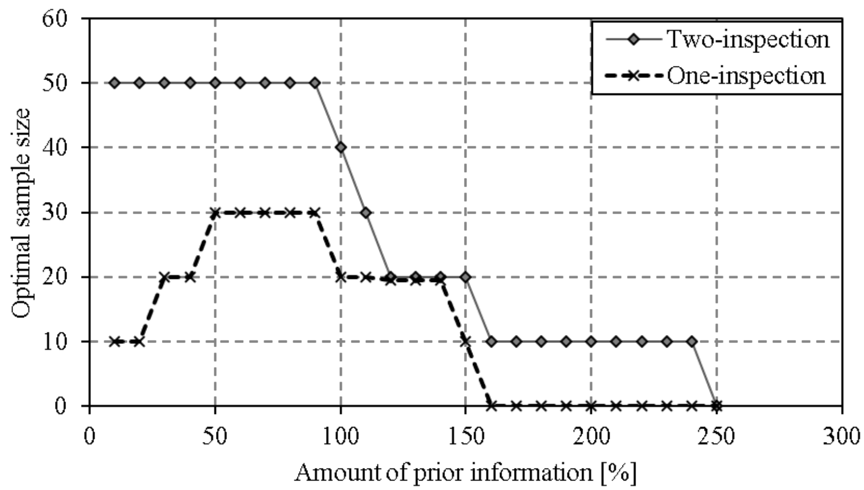

Figure 7.6: Optimal sample size with different amounts of prior information ($C_I = 3$, $C_P = 10$, and $C_F = 100$)

inspection cost is not low enough, the operators need to calculate $ENGS$ for the overall components and select the appropriate sample size for each inspection outage.

7.6 Generalized Multiple-Inspection Problem

This section generalizes the VoI-based SSD method to multiple-inspection problems as dynamic programming equations.

7.6.1 Maintenance Stage

For each action, a, the cost is expected to rise between t_i and t_{i+1}, and is determined by four parameters: $T_{insp}^{(i)}$, $S_{insp}^{(i)}$, x_{last}, and Δt_{no}. $T_{insp}^{(i)}$ denotes the total inspection interval and is defined as $T_{insp}^{(i)} = \sum_{j=1}^{N} t_{j,insp}$; $S_{insp}^{(i)}$ is the accumulation of all inspected deteriorations and is defined as $S_{insp}^{(i)} = \sum_{j=1}^{N} x_{j,insp}$; x_{last} is the latest observed degradation level of a single component; and Δt_{no}

denotes the time difference between t_{i+1} and the last inspection time, t_{last}. Note that the first two parameters, $T_{insp}^{(i)}$ and $S_{insp}^{(i)}$, update the estimation for the parameter of the degradation process, μ, and the last two, x_{last} and Δt_{no}, affect the probability of failure in an inspection/maintenance interval. Let us set parameters for parameter updating, $\Theta^{(i)} = \left(T_{insp}^{(i)}, S_{insp}^{(i)}\right)$, and parameters for the inspection outcome of a component, $\theta_j^{(i)} = \left(x_{j,last}^{(i)}, \Delta t_{j,no}^{(i)}\right)$. Suppose the optimal expected cost for each component at t_{i+1} is given as $C_j^{(i+1)}\left(a^{(i+1)o} \mid e^{(i+1)o}, n^{(i+1)o}, \Theta^{(i+1)}, \theta_j^{(i+1)}\right)$. The optimal expected cost is calculated as follows:

$$\begin{aligned}
& C_j^{(i)}\left(a^{(i)o} \mid e^{(i)}, n^{(i)}, \Theta^{(i)}, \theta_j^{(i)}\right) \\
= & \min_{a^{(i)}} \left\{ C_P, C_F \mathbb{E}_{M|\Theta^{(i)}}\left[1 - F_{X(\Delta t_{j,no}^{(i+1)})}(\rho_F - x_{j,last}^{(i)})\right] \right. \\
& \left. + C_j^{(i+1)}\left(a^{(i+1)o} \mid e^{(i+1)o}, n^{(i+1)o}, \Theta^{(i+1)}, \theta_j^{(i+1)}\right) \mathbb{E}_{M|\Theta^{(i)}}\left[F_{X(\Delta t_{j,no}^{(i+1)})}(\rho_F - x_{j,last}^{(i)})\right] \right\}.
\end{aligned} \tag{7.58}$$

7.6.2 Inspection Stage

In the inspection stage, we accumulate the expected cost for individual components and obtain the optimal sample size, $n^{(i)}$. Inspection outcomes update $\Theta^{(i-1)}$ and $\theta_j^{(i-1)}$ to $\Theta^{(i)}$ and $\theta_j^{(i)}$, respectively; if the jth component is not inspected, $\theta_j^{(i)} = \theta_j^{(i-1)}$, and if no component is inspected, $\Theta^{(i)} = \Theta^{(i-1)}$. The expected prior cost is

$$\begin{aligned}
C_{prior}^{(i)} = & \min_{a^{(i)}} \left\{ C_P, C_F \mathbb{E}_{M|\Theta^{(i-1)}}\left[1 - F_{X(\Delta t_{j,no}^{(i+1)})}(\rho_F - x_{j,last}^{(i-1)})\right] \right. \\
& \left. + C_j^{(i+1)}\left(a^{(i+1)o} \mid e^{(i+1)o}, n^{(i+1)o}, n^{(i)} = 0, \Theta^{(i+1)}, \theta_j^{(i+1)}\right) \mathbb{E}_{M|\Theta^{(i-1)}}\left[F_{X(\Delta t_{j,no}^{(i+1)})}(\rho_F - x_{j,last}^{(i-1)})\right] \right\}.
\end{aligned} \tag{7.59}$$

The expected pre-posterior cost for an un-inspected component is

$$C_{un}^{(i)} = \mathbb{E}_{\theta^{(i)},\Theta^{(i)},M|\Theta^{(i-1)}} \left[C_j^{(i)} \left(a^{(i)o} \mid e_2^{(i)}, n^{(i)}, \Theta^{(i)}, \theta_j^{(i)} \right) \right] \tag{7.60}$$

The expected pre-posterior cost for an inspected component is

$$C_{insp}^{(i)} = C_I + \mathbb{E}_{\theta^{(i)},\Theta^{(i)},M|\Theta^{(i-1)}} \left[C_j^{(i)} \left(a^{(i)o} \mid e_1^{(i)}, n^{(i)}, \Theta^{(i)}, \theta_j^{(i)} \right) \right] \tag{7.61}$$

Thus, the $ENGS^{(i)}$ is derived as follows:

$$ENGS^{(i)}(n^{(i)}) \quad = \quad NC_{prior}^{(i)} - (N - n^{(i)})C_{un}^{(i)} - n^{(i)}C_{insp}^{(i)}. \tag{7.62}$$

Finally, the optimal sample size, $n^{(i)o}$, is determined as follows:

$$n^{(i)o} \quad = \quad \arg\max_{n^{(i)}} ENGS^{(i)}(n^{(i)}) \tag{7.63}$$

Note that the optimal expected cost for each component at t_i can be derived as $C_j^{(i)} \left(a^{(i)o} \mid e^{(i)o}, n^{(i)o}, \Theta^{(i)}, \theta_j^{(i)} \right)$ with the obtained $n^{(i)o}$. Until the time horizon becomes t_1, the backward process is repeated. Once the process reaches t_1, the most cost-effective inspection/maintenance policy is proposed over the whole decision-making problem.

7.7 Summary

A statistical SSD method for system-level condition-based maintenance has been developed for a two-inspection problem. This study has introduced how the VoI approach can be applied to a multiple-inspection problem. The two-inspection problem is understood in the context of a general sequential decision problem with dynamic programming that indicates how to solve the problem with backward induction. The analysis process has been formulated in a two-inspection problem, and used to derive the definition of the *ENGS*. Numerical analysis with the data from a real operating situation illustrates how the results support operators in their decision-making. Based on the method for a two-inspection problem, the VoI-based SSD method is theoretically generalized to a multiple-inspection problem.

Chapter 8

Conclusions and Recommendations

8.1 Conclusions

Based on the value of information (VoI) concept, this thesis has formulated and developed statistical sample size determination (SSD) methods for the maintenance of engineering components that follow stochastic degradation process models. These methods provide an inspection policy in a context of condition-based maintenance by defining maintenance problems as sequential decision-making problems. The VoI-based methods are applied to two degradation models: the random rate model and the gamma process model.

Various existing SSD methods are summarized, and their advantages and disadvantages are discussed. Based on the classification, existing standards for engineering-components are then categorized, and the rationale behind each standard is described. The categorization reveals that the existing standards for SSD rely on deterministic criteria that are not theoretically supported, or the criteria obtained through traditional SSD method: the hypothesis-testing approach. This study compares the VoI-based method with the hypothesis-testing method by formulating a situation in which traditional hypothesis-testing approaches have been used as a Bayesian sequential-decision making problem. The stated problem uses a binomial component's state, in what can be classified as a special case of the random rate model. The superiority of the VoI-based approach is highlighted through conceptual, theoretical, and numerical analyses.

The VoI-based SSD method is applied to a degradation model, the random rate model. Through formulating the VoI-based SSD method and demonstrating the model in a numerical example, the advantages and the limitations of the random rate model are discussed.

Adding the gamma process to extend the model to condition-based maintenances enables the VoI-based SSD method to be used in realistic maintenance and inspection problems. The model mathematically describes how temporal and parameter uncertainties of the degradation process affect VoI-based analysis. Because of the additivity characteristics of the gamma process, the VoI-based SSD method can be formulated simply. To reduce computational costs, two computational calculation techniques are introduced and applied to the proposed method. Their costs and accuracies are compared, and the more effective method, the probability mass function (PMF) approach, is used in the numerical analysis. An application to a real degrading system demonstrates the effectiveness of the approach. The sensitivity of each parameter and the contributions of the reducing parameter and temporal uncertainties are analysed and discussed.

For more-generalized condition-based maintenance applications, the thesis further extends the model to a multiple-inspection problem by generalizing the model as a dynamic programming problem. A problem with data from a real operating system is numerically analysed with a two-inspection problem and shows how the first inspection subsequently affects maintenance and second-inspection decisions. An optimal inspection and maintenance policy is proposed and illustrated through the example.

The major insights found for sample size determination strategies are as follows:

- VoI-based SSD has advantages that can be shown qualitatively and quantitatively;

- In most cases, a re-inspection has lower priority than a first-time-inspection;
- The optimal sample size is sensitive against parameters of cost and prior distribution;
- Reducing temporal uncertainty contributes to increasing *ENGS*, through the benefit from inspected components;
- Reducing parameter uncertainty enhances *ENGS*, mostly from better estimations for un-inspected components through parameter updating;
- The chance of a second inspection enhances the value of the first inspection.

8.2 Recommendations for Further Research

This study has several opportunities for expansion that warrant further research. First, latent failure mode should be considered, to make the model applicable to several other fields, where not only physical failures but also standard violations are important in the maintenance of operating systems. Under stringent standards and regulations of operation, operators are actually concerned about the penalty of finding latent failure in components, as this failure may raise concerns about critical system failure, even though such components may still continue to work safely.

Second, the generalized model for multiple-inspection problems requires computationally more-efficient techniques for numerical calculations than the proposed method. The maximum number of time horizons of the current model has been roughly estimated as four, if the SSD method for the four-inspection problem could be modelled as a simple extension of the current model for a two-inspection problem. The challenge in solving the problem is to reduce the computational cost of calculating high-dimensional integrations.

Third, cost functions should be investigated for more-realistic applications. In this study, the costs of inspection, replacement, and failure are treated as fixed values for any situation; however, these costs can be functions of sample size, population size, or time. Moreover, if a system failure may increase health risks to workers or residents around a site, human lives need to be estimated in monetary terms to consider the cost-risk trade-off. The well-known value for human life is the "value of statistical life (VSL)," which is calculated from people's willingness to pay (WTP) to avoid risk.

References

Abdel-Hameed, M. (1975). A gamma wear process. *IEEE Transactions on Reliability*, R-24(2):152–153.

Adcock, C. (1988). A bayesian approach to calculating sample sizes. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 37(4/5):433–439.

Adcock, C. (1997). Sample size determination: a review. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 46(2):261–283.

Ahmadi, R. (2016). An optimal replacement policy for complex multi-component systems. *International Journal of Production Research*, 54(17):5303–5316.

American Society for Testing and Materials (2009). ASTM E122-09: *Standard practice for calculating sample size to estimate, with specified precision, the average for a characteristic of a lot or process.* Technical report, West Conshohocken, PA, United States.

American Society for Testing and Materials (2015). ASTM F302-09: *Standard practice for field sampling of aerospace fluids in containers.* Technical report, West Conshohocken, PA, United States.

American Society for Testing and Materials (2016). ASTM B602-88: *Standard test method for attribute sampling of metallic and inorganic coatings.* Technical report, West Conshohocken, PA, United States.

American Society of Civil Engineers (2016). Failure to act: closing the infrastructure investment gap for america's economic future. Technical report, Boston, MA, United States.

Bellman, R. (1957). *Dynamic programming.* Princeton University Press, Princeton, NJ, United States.

Bellman, R. and Lee, E. (1984). History and development of dynamic programming. *Control Systems Magazine, IEEE*, 4(4):24–28.

Benjamin, J. R. and Cornell, C. A. (1970). *Probability, statistics, and decision for civil engineers.* McGraw-Hill, New York, NY, United States.

Bensi, M. T. (2010). *A Bayesian network methodology for infrastructure seismic risk assessment and decision support.* PhD thesis, University of California, Berkeley.

Blackwell, D. and Girshick, M. A. (1954). *Theory of games and statistical decisions.* Wiley, New York, NY, United States.

Boardman, A. E. (2006). *Cost-benefit analysis: concepts and practice.* Pearson/Prentice Hall, Upper Saddle River, NJ, United States, third edition.

Bumblauskas, D. (2015). A Markov decision process model case for optimal maintenance of serially dependent power system components. *Journal of Quality in Maintenance Engineering*, 21(3):271–293.

Chen, N., Ye, Z. S., Xiang, Y., and Zhang, L. (2015). Condition-based maintenance using the inverse Gaussian degradation model. *European Journal of Operational Research*, 243(1):190–199.

Chernoff, H. and Moses, L. E. (1959). *Elementary decision theory.* Wiley, New York, NY, United States.

Cinlar, E., Bazant, Z., and Osman, E. (1977). Stochastic process for extrapolating concrete creep. *Journal of Engineering Mechanics - ASCE*, 103(6):1069–1088.

Claxton, K. and Sculpher, M. (2006). Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmaco Economics*, 24(11):1055–1068.

Cochran G., W. (1977). *Sampling techniques*. John Wiley & Sons, Inc, New York, NY, United States, third edit edition.

DeGroot, M. H. (1970). *Optimal statistical decisions*. Mcgraw-Hill, New York, NY, United States.

Desu, M. and Raghavarao, D. (1990). Sample size methodology. *Statistical modeling and decision science Show all parts in this series*.

Dooley, R. B. and Chexal, V. K. (2000). Flow-accelerated corrosion of pressure vessels in fossil plants. *International Journal of Pressure Vessels and Piping*, 77(2-3):85–90.

Eisenhart, C., Hastay, M. W., and Wallis, W. A. (1947). Techniques of statistical analysis. *New York*, page 98.

Electric Power Research Institute (1999). TR-017218-R1: *Guideline for sampling in the commercial-grade item acceptance process*. Technical Report Report TR-017218-R1, Palo Alto, CA, United States.

Faber, M. H. and Sorensen, J. D. (2002). Indicators for inspection and maintenance planning of concrete structures. *Structural Safety*, 24(2-4):377–396.

Faber, M. H., Sorensen, J. D., Tychsen, J., and Straub, D. (2005). Field implementation of RBI for jacket structures. *Journal of Offshore Mechanics and Arctic Engineering*, 127(3):220.

Faddoul, R., Raphael, W., and Chateauneuf, A. (2011). A generalised partially observable Markov decision process updated by decision trees for maintenance optimisation. *Structure and Infrastructure Engineering*, 7(10):783–796.

Ferguson, T. S. (1967). *Mathematical statistics: a decision theoretic approach.* Academic Press, New York, NY, United States.

Frangopol, D. M. and Soliman, M. (2016). Life-cycle of structural systems: recent achievements and future directions. *Structure and Infrastructure Engineering*, 12(1):1–20.

Garland, W. J., editor (2014). *The essential CANDU, A textbook on the CANDU nuclear power plant technology.* University Network of Excellence in Nuclear Engineering (UNENE). Retrieved from https://www.unene.ca/education/candu-textbook on Jan 24, 2018.

Guenther, W. C. (1981). Sample size formulas for normal theory t tests. *The American Statistician*, 35(4):243–244.

Hadley, G. (1967). *Introduction to probability and statistical decision theory.* Holden-Day, San Francisco, CA, United States.

Higo, E. and Pandey, M. D. (2015). A required level of enhancing life safety derived from the cost for substituting nuclear energy in Japan. *IDRiM Journal*, 5(2):153–166.

Higo, E. and Pandey, M. D. (2016). Sample size determination in engineering component inspection, a comparison. *ASME 2016 Pressure Vessels and Piping Conference*, pages 1–9.

Hong, H. (2000). Assessment of reliability of aging reinforced concrete structures. *Journal of Structural Engineering*, 126(12):1458–1465.

Jalal, H., Goldhaber-Fiebert, J. D., and Kuntz, K. M. (2015). Computing expected value of partial sample information from probabilistic sensitivity analysis using linear regression metamodeling. *Medical Decision Making*, 35(5):584–595.

Juran, J. M., Gryna, F. M., and Bingham, R. S. (1974). *Quality control handbook*. McGrawHill Book Company, third edition.

Kallen, M. J. and van Noortwijk, J. M. (2005). Optimal maintenance decisions under imperfect inspection. *Reliability Engineering and System Safety*, 90(2-3):177–185.

Kele, B., Tekin, S., and Bakir, N. O. (2017). Maintenance policies for a deteriorating system subject to non-self-announcing failures. *IEEE Transactions on Reliability*, 66(1):219–232.

Khalifa, M., Khan, F., and Haddara, M. (2012). Bayesian sample size determination for inspection of general corrosion of process components. *Journal of Loss Prevention in the Process Industries*, 25(1):218–223.

Konakli, K., Sudret, B., and Faber, M. H. (2016). Numerical investigations into the value of information in lifecycle analysis of structural systems. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 2(3):B4015007.

Kraemer, H. and Thiemann, S. (1988). *How many subjects? : Statistical power analysis in research*. Sage Publications, Newbury Park, CA, United States.

Krause, A. (2008). *Optimizing sensing: Theory and applications*. PhD thesis, Carnegie Mellon University.

Kullback, S. (1968). *Information theory and statistics*. Dover Publications, New York, NY, United States.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.

Lawless, J. and Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis*, 10(3):213–227.

Lin, S., Zhang, A., and Feng, D. (2016). Maintenance decision-making model based on POMDP for traction power supply equipment and its application. In *Proceedings of 2016 prognostics and system health management conference, PHM-Chengdu*, pages 1–6.

Lindley, D. V. (1997). The choice of sample size. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 46(2):129–138.

Madsen, H. O. (1997). *Stochastic modeling of fatigue crack growth and inspection.* In Soares, C. G. (Ed.)*, Probabilistic methods for structural design*, pages 59–83. Springer.

Madsen, H. O., Krenk, S., and Lind, N. C. (1986). *Methods of structural safety.* Prentice-Hall, Englewood Cliffs, NJ.

McGrayne, S. B. (2011). *The theory that would not die: how Bayes' rule cracked the enigma code, hunted down Russian submarines, & emerged triumphant from two centuries of controversy.* Yale University Press.

Memarzadeh, M. and Pozzi, M. (2016). Integrated inspection scheduling and maintenance planning for infrastructure systems. *Computer-Aided Civil and Infrastructure Engineering*, 31(6):403–415.

Miller, T. R. (2000). Variations between countries in the evalues of statistical life. *Journal of Transport Economics and Policy*, 34(2):169–188.

Mrozek, J. R. and Taylor, L. O. (2002). What determines the value of life?: A meta-analysis. *Journal of Policy Analysis and Management*, 21(2):253–270.

Nathwani, J. S., Lind, N. C., and Pandey, M. D. (1997). *Affordable safety by choice: the life quality method*. Institute for Risk Research, University of Waterloo, Waterloo, ON, Canada.

Nathwani, J. S., Lind, N. C., and Pandey, M. D. (2009). *Engineering decisions for life quality: How safe is safe enough?* Springer, London, United Kingdom.

National Standards of Canada (2014). N285.4-14: *Periodic inspection of CANDU nuclear power plant components*. Technical report, Mississauga, Ontario, Canada.

Nguyen, T. P. K., Yeung, T. G., and Castanier, B. (2013). Optimal maintenance and replacement decisions under technological change with consideration of spare parts inventories. *International Journal of Production Economics*, 143(2):472–477.

Oakley, J. E., Brennan, A., Tappenden, P., and Chilcott, J. (2010). Simulation sample sizes for Monte Carlo partial EVPI calculations. *Journal of Health Economics*, 29(3):468–477.

Owen, D. B. (1962). *Handbook of statistical tables*, volume 117. Addison-Wesley Reading, MA.

Pandey, M. (1998). Pobabilistic models for condition assessment of oil and gas pipelines. *NDT & E International*, 31(5):349–358.

Pandey, M. D., Nathwani, J. S., and Lind, N. C. (2006). The derivation and calibration of the life-quality index (LQI) from economic principles. *Structural Safety*, 28:341–360.

Pandey, M. D., Yuan, X. X., and van Noortwijk, J. M. (2009). The influence of temporal uncertainty of deterioration on life-cycle management of structures. *Structure and Infrastructure Engineering*, 5(2):145–156.

Papakonstantinou, K. G. and Memarzadeh, M. (2017). Optimal maintenance and inspection planning for struc- tural components under mixed and partial observability. In *12th Int. Conf. on Structural Safety and Reliability*, pages 2280–2289, Vienna, Austria.

Papakonstantinou, K. G. and Shinozuka, M. (2014a). Optimum inspection and maintenance policies for corroded structures using partially observable Markov decision processes and stochastic, physically based models. *Probabilistic Engineering Mechanics*, 37:93–108.

Papakonstantinou, K. G. and Shinozuka, M. (2014b). Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part I: Theory. *Reliability Engineering and System Safety*, 130.

Papakonstantinou, K. G. and Shinozuka, M. (2014c). Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation. *Reliability Engineering and System Safety*, 130:214–224.

Peng, W., Li, Y.-f., Yang, Y.-j., Mi, J., and Huang, H.-z. (2017). Bayesian degradation analysis with inverseGaussian process models under time-varying degradation rates. *IEEE Transactions on Reliability*, 66(1):84–96.

Pham-Gia, T. and Turkkan, N. (1992). Sample size determination in bayesian analysis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 41(4):389–397.

Pozzi, M. and Der Kiureghian, A. (2011). Assessing the value of information for long-term structural health monitoring. In *Health Monitoring of Structural and Biological Systems 2011*, volume 7984.

Pozzi, M. and Der Kiureghian, A. (2012). Assessing the value of alternative bridge health monitoring systems. In *6th International Conference on Bridge Maintenance, Safety and Management, IABMAS. Como, Italy.*

Pratt, J. W., Raiffa, H., and Schlaifer, R. (1995). *Introduction to statistical decision theory.* MIT press.

Qin, H., Zhang, S., and Zhou, W. (2013). Inverse Gaussian process-based corrosion growth modeling and its application in the reliability analysis for energy pipelines. *Frontiers of Structural and Civil Engineering*, 7(3):276–287.

Raiffa, H. and Schlaifer, R. (1961). *Applied statistical decision theory.* Harvard University, Boston, MA, United States.

Sandiford, P. J. (1960). A new binomial approximation for use in sampling from finite populations. *Journal of the American Statistical Association*, 55(292):718–722.

Schlaifer, R. (1959). *Probability and statistics for business decisions; an introduction to managerial economics under uncertainty.* McGRAW-HILL BOOK COMPANY, INC.

Schöbi, R. and Chatzi, E. N. (2016). Maintenance planning using continuous-state partially observable Markov decision processes and non-linear action models. *Structure and Infrastructure Engineering*, 12(8):977–994.

Schouten, H. J. (1999). Sample size formula with a continuous outcome for unequal group sizes and unequal variances. *Statistics in medicine*, 18(1):87–91.

Srinivasan, R. and Parlikad, A. K. N. (2014). Semi Markov decision process with partial information for optimum maintenance decisions. *IEEE Transactions on Reliability*, 63(4):891–898.

Stein, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *The Annals of Mathematical Statistics*, 16(3):243–258.

Steuten, L., van de Wetering, G., Groothuis-Oudshoorn, K., and Retèl, V. (2013). A systematic and critical review of the evolving methods and applications of value of information in academia and practice. *Pharmacoeconomics*, 31(1):25–48.

Stewart, M. and Rosowsky, D. (1998). Time-dependent reliability of deteriorating reinforced concrete bridge decks. *Structural Safety*, 20(1):91–109.

Straub, D. (2004). *Generic approaches to risk based inspection planning for steel structures*. PhD thesis, Swiss Federal Institute of Technology.

Straub, D. (2014). Value of information analysis with structural reliability methods. *Structural Safety*, 49:75–85.

Straub, D. and Faber, M. H. (2004a). On the relation between inspection quantity and quality. *Journal of Nondestructive Testing*, 9(7):1–14.

Straub, D. and Faber, M. H. (2004b). System effects in generic risk based inspection planning. *Journal of Offshore Mechanics and Arctic Engineering*, 126(3):265–271.

Straub, D. and Faber, M. H. (2005). Risk based inspection planning for structural systems. *Structural Safety*, 27(4):335–355.

U.S. Department of Defence (1957). MIL-STD-414: *Sampling procedures and tables for inspection by variables for percent defectives*. Technical report, Washington, D.C., United States.

U.S. Department of Defence (1989). MIL-STD-105E:textitSampling procedures and tables for inspection by attributes. Technical report, Washington, D.C., United States.

U.S. Environmental Protection Agency (2006). EPA QA/G-4: *Guidance on systematic planning using the data quality objectives process.* Technical report, Washington, D.C., United States.

U.S. Environmental Protection Agency (2011). The benefits and sosts of the clean air act from 1990 to 2020: Final report. Technical report.

U.S. Nuclear Regulatory Commission (1998). NUREG-1505: *A nonparametric statistical methodology for the design and analysis of final status decommissioning surveys*. Technical report, New York, NY, United States.

U.S. Nuclear Regulatory Commission (2011). NUREG-1475: *Applying statistics*. Technical report, Washington, D.C., United States.

van der Weide, J. A. M., Pandey, M. D., and van Noortwijk, J. M. (2010). Discounted cost model for condition-based maintenance optimization. *Reliability Engineering and System Safety*, 95(3):236–246.

van Noortwijk, J. M. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering and System Safety*, 94(1):2–21.

van Noortwijk, J. M., Cooke, R. M., and Kok, M. (1995). A Bayesian failure model based on isotropic deterioration. *European Journal of Operation*, 82:270–282.

van Noortwijk, J. M. and Klatter, H. E. (1999). Optimal inspection decisions for the block mats of the Eastern-Scheldt barrier. *Reliability Engineering and System Safety*, 65(3):203–211.

van Noortwijk, J. M. and Peerbolte, E. B. (2000). Optimal sand nourishment decisions. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 126(1):30–38.

Verbeke, G. and Molenberghs, G. (2009). *Linear mixed models for longitudinal data.* Springer Science & Business Media.

Viscusi, W. K. (1992). *Fatal tradeoffs : public and private responsibilities for risk.* Oxford University Press, New York, NY, United States.

Viscusi, W. K. and Aldy, J. E. (2003). *The value of a statistical life: a critical review of market estimates throughout the world*, volume 9487.

von Neumann, J. and Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton University Press, Princeton, NJ, United States, second edition.

Wald, A. (1947). *Sequential analysis.* Wiley, New York, NY, United States.

Wald, A. (1950). *Statistical decision functions.* Wiley, New York, NY, United States.

Wang, X. (2010). Wiener processes with random effects for degradation data. *Journal of Multivariate Analysis*, 101(2):340–351.

Wang, X. and Xu, D. (2010). An inverse Gaussian process model for degradation data. *Technometrics*, 52(2):188–197.

Weiss, L. (1961). *Statistical decision theory.* McGraw-Hill, New York, NY, United States.

Xing, Y. Y., Jiang, P., and Cheng, Z. J. (2016). The determination method on products sample size under the condition of Bayesian sequential testing. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 1635–1639.

Ye, Z. S. and Chen, N. (2014). The inverse Gaussian process as a degradation model. *Technometrics*, 56(3):302–311.

Ye, Z.-s., Chen, N., and Shen, Y. (2015). A new class of Wiener process models for degradation analysis. *Reliability Engineering and System Safety*, 139:58–67.

Ye, Z. S., Wang, Y., Tsui, K. L., and Pecht, M. (2013). Degradation data analysis using wiener processes with measurement errors. *IEEE Transactions on Reliability*, 62(4):772–780.

Yoshida, I. (2015). Parameter study on optimal sampling planning based on value of information. In *12th International Conference on Applications of Statistics and Probability in Civil Engineering (Vancouver, Canada, July 12-15, 2015)*.

Yuan, X.-X. (2007). *Stochastic modeling of deterioration in nuclear power plant components*. PhD thesis, University of Waterloo.

Yuan, X.-X., Pandey, M. D., and Bickel, G. (2006). A probabilistic degradation model for the estimation of the remaining life distribution of feeders. In *Annual conference of the Canadian Nuclear Society*, Toronto, ON, Canada.

Yuan, X.-X., Pandey, M. D., and Bickel, G. A. (2008). A probabilistic model of wall thinning in CANDU feeders due to flow-accelerated corrosion. *Nuclear Engineering and Design*, 238:16–24.

Zhang, M. and Revie, M. (2017). Continuous-observation partially observable semi-Markov decision processes for machine maintenance. *IEEE Transactions on Reliability*, 66(1):202–218.

APPENDICES

Appendix A

Derivation of the *EVSI* and the *ENGS*: Chapter 3

This section shows how to derive $EVSI$ and $ENGS$ in Equations (3.16) and (3.18), respectively. The calculation for expected pre-posterior cost, shown in Equation (3.14), becomes

$$\begin{aligned}
& \mathbb{E}_W\left[\mathbb{E}_{X|W}\left[C(a^o(w), X \mid n, w)\right]\right] \\
= \ & \mathbb{E}_W\left[\min_a \left\{\mathbb{E}_{X|W}\left[C(a, X \mid n, w)\right]\right\}\right] \\
= \ & \sum_{w=0}^{h_n}\left[(N-n)\bar{x}''(n,w)C_F + wC_P\right]\cdot f_W(w \mid n) \\
& + \sum_{w=h_n+1}^{n}\left[(N-n)x_b(C_F - C_P) + (N-n)\bar{x}''(n,w)C_P + wC_P\right]\cdot f_W(w \mid n) \qquad \text{(A.1)}
\end{aligned}$$

Thus, the *EVSI* is derived as follows:

$EVSI(n) = \min_a[\mathbb{E}_X\left[C(a,X)\right]] - \mathbb{E}_W\left[\min_a\left\{\mathbb{E}_{X|W}\left[C(a,X)\right]\right\}\right]$

$$= \begin{cases} Nx_b(C_F - C_P) + N\bar{x}C_P - \bar{w}C_P - \sum_{w=0}^{h_n} [(N-n)\bar{x}''(n,w)C_F] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\ \quad - \sum_{w=h_n+1}^{n} [(N-n)x_b(C_F - C_P) + (N-n)\bar{x}''(n,w)C_P] \cdot f_W(w \mid n) & \\ N\bar{x}C_F - \bar{w}C_P - \sum_{w=0}^{h_n} [(N-n)\bar{x}''(n,w)C_F] \cdot f_W(w \mid n) & \text{if } \bar{x} \le x_b \\ \quad - \sum_{w=h_n+1}^{n} [(N-n)x_b(C_F - C_P) + (N-n)\bar{x}''(n,w)C_P] \cdot f_W(w \mid n), & \end{cases}$$

$$= \begin{cases} \left[N - \sum_{w=h_n+1}^{n}(N-n)f_W(w \mid n)\right] x_b(C_F - C_P) & \\ \quad +(N-n)\bar{x}C_P - \sum_{w=0}^{h_n} [(N-n)\bar{x}''(n,w)C_F] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\ \quad - \sum_{w=h_n+1}^{n} [(N-n)\bar{x}''(n,w)C_P] \cdot f_W(w \mid n) & \\ - \sum_{w=h+1}^{n} [(N-n)x_b(C_F - C_P)] f_W(w \mid n) & \\ \quad +(N-n)\bar{x}C_F + n\bar{x}(C_F - C_P) & \\ \quad - \sum_{w=0}^{h_n} [(N-n)\bar{x}''(n,w)C_F] \cdot f_W(w \mid n) & \text{if } \bar{x} \le x_b \\ \quad - \sum_{w=h_n+1}^{n} [(N-n)\bar{x}''(n,w)C_P] \cdot f_W(w \mid n), & \end{cases}$$

$$= \begin{cases} \left[N - \sum_{w=h_n+1}^{n}(N-n)f_W(w \mid n)\right] x_b(C_F - C_P) & \\ \quad - \sum_{w=0}^{h_n} [(N-n)\bar{x}''(n,w)(C_F - C_P)] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\ - \sum_{w=h_n+1}^{n}(N-n)x_b(C_F - C_P)f_W(w \mid n) + n\bar{x}(C_F - C_P) & \\ \quad + \sum_{w=h_n+1}^{n} [(N-n)\bar{x}''(n,w)(C_F - C_P)] \cdot f_W(w \mid n) & \text{if } \bar{x} \le x_b, \end{cases} \tag{A.2}$$

where we use

$$\bar{w} = n\bar{x} \tag{A.3}$$

$$\bar{x} = \sum_{i=1}^{20} x_i f_X(x_i) = \sum_{i=1}^{20} x_i \sum_{w=0}^{n} f_{X,W}(x_i, w \mid n)$$

$$= \sum_{w=0}^{n} \sum_{i=1}^{20} x_i f_{X|W}(x_i \mid n, w) f_W(w \mid n). \tag{A.4}$$

The *ENGS* in Equation (3.18) is derived as follows:

$$
\begin{aligned}
& ENGS(n) \\
= \; & EVSI(n) - nC_I = EVSI(n) - nx_b(C_F - C_P) \\
= \; & \begin{cases}
\left[1 - \sum_{w=h_n+1}^{n} f_W(w \mid n)\right](N-n)x_b(C_F - C_P) & \\
\quad - \sum_{w=0}^{h_n} [(N-n)\bar{x}''(n,w)(C_F - C_P)] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\
- \sum_{w=h_n+1}^{n} (N-n)x_b(C_F - C_P) f_W(w \mid n) & \\
\quad + n(\bar{x} - x_b)(C_F - C_P) & \\
\quad + \sum_{w=h_n+1}^{n} [(N-n)\bar{x}''(n,w)(C_F - C_P)] \cdot f_W(w \mid n) & \text{if } \bar{x} \le x_b,
\end{cases} \\
= \; & \begin{cases}
(N-n)(C_F - C_P) \sum_{w=0}^{h_n} [x_b - \bar{x}''(n,w)] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\
(N-n)(C_F - C_P) \sum_{w=h_n+1}^{n} [\bar{x}''(n,w) - x_b] \cdot f_W(w \mid n) & \\
\quad + n(C_F - C_P)(\bar{x} - x_b) & \text{if } \bar{x} \le x_b,
\end{cases} \\
= \; & \begin{cases}
(N-n)C_I \sum_{w=0}^{h_n} [1 - \bar{x}''(n,w)/x_b] \cdot f_W(w \mid n) & \text{if } \bar{x} > x_b \\
(N-n)C_I \sum_{w=h_n+1}^{n} [\bar{x}''(n,w)/x_b - 1] \cdot f_W(w \mid n) - nC_I(1 - \bar{x}/x_b) & \text{if } \bar{x} \le x_b,
\end{cases}
\end{aligned}
\tag{A.5}
$$

where we use

$$
\bar{x} - \sum_{w=h_n+1}^{n} \bar{x}''(n,w) f_W(w \mid n) = \sum_{w=0}^{h_n} \bar{x}''(n,w) f_W(w \mid n) \tag{A.6}
$$

$$
1 - \sum_{w=h_n+1}^{n} f_W(w \mid n) = \sum_{w=0}^{h_n} f_W(w \mid n). \tag{A.7}
$$

Appendix B

Posterior Distribution: Chapter 6

The likelihood function of μ is

$$\begin{aligned} L(\mu \mid x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)}) &= \prod_{j=1}^{n} f_X(x_1^{(j)} \mid \mu) \\ &= \frac{\left(\frac{1}{\mu\nu^2}\right)^{\frac{nt_1}{\nu^2}}}{\left(\Gamma\left(\frac{t_1}{\nu^2}\right)\right)^n} \prod_{j=1}^{n} x_1^{(j)\,\frac{t_1}{\nu^2}-1} exp\left(-\frac{\sum_{j=1}^{n} x_1^{(j)}}{\mu\nu^2}\right). \end{aligned} \tag{B.1}$$

The posterior distribution for X is calculated using Bayes' rule.

$$\begin{aligned} f_M(\mu \mid x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)}) &= \frac{L(\mu \mid x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)}) f_M(\mu)}{\int_0^\infty L(\mu \mid x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)}) f_M(\mu) d\mu} \\ &= \frac{K\left(\frac{1}{\mu}\right)^{\alpha+\frac{nt_1}{\nu^2}+1} exp\left(-\frac{1}{\mu}\left(\beta + \frac{\sum_{j=1}^{n} x_1^{(j)}}{\nu^2}\right)\right)}{K\int_0^\infty \left(\frac{1}{\mu}\right)^{\alpha+\frac{nt_1}{\nu^2}+1} exp\left(-\frac{1}{\mu}\left(\beta + \frac{\sum_{j=1}^{n} x_1^{(j)}}{\nu^2}\right)\right) d\mu} \end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{\mu}\right)^{\alpha+\frac{nt_1}{\nu^2}+1} exp\left(-\frac{1}{\mu}\left(\beta+\frac{\sum_{j=1}^{n} x_1^{(j)}}{\nu^2}\right)\right) \\
&= Iga\left(\mu;\alpha+\frac{nt_1}{\nu^2},\beta+\frac{\Sigma_{j=1}^{n}x_1^{(j)}}{\nu^2}\right), \qquad \text{(B.2)}
\end{aligned}
$$

where K is a normalization factor of the probability distribution and is not a function of μ, as it is

$$
K = \frac{\left(\frac{1}{\nu^2}\right)^{\frac{nt_1}{\nu^2}}\beta^{\alpha}}{\left(\Gamma\left(\frac{t_1}{\nu^2}\right)\right)^{n}\Gamma(\alpha)} \prod_{j=1}^{n} x_1^{(j)\,\frac{t_1}{\nu^2}-1}. \qquad \text{(B.3)}
$$

Because $s_n = \Sigma_{j=1}^{n} x_1^{(j)}$ is the sufficient statistics, the posterior distribution can be simplified as

$$
f_M(\mu \mid s_n) \;=\; Iga\left(\mu;\alpha+\frac{nt_1}{\nu^2},\beta+\frac{s_n}{\nu^2}\right). \qquad \text{(B.4)}
$$

Thus, operators only need to consider the sum of the inspected n components' conditions, s_n, for updating the PDF of the unknown parameter, μ.

Appendix C

Indices for Difference between Two Distributions

This appendix introduces two indices, the histogram intersection and the Kulback-Leibler divergence, which can represent difference between two distributions, especially for discrete distributions.

C.1 Histogram Intersection

The simplest method to illustrate the similarity between two distributions is the histogram intersection (HI). Given histograms of the two distributions (or two PMFs), the HI is calculated as follows:

$$HI = \sum_{i} \min(Q_1(i), Q_2(i)), \tag{C.1}$$

where $Q_j(i)$ for $j = 1, 2$ is the probability mass for the ith interval. The index can take a value between zero, which means no similarity, and one, which indicate that the two distribution is the same.

C.2 Kallback-Leibler Divergence

The Kullback-Leibler (KL) divergence was first introduced by Kullback and Leibler (1951) as a directed divergence between two distributions and has been used to evaluate how much one distribution is different from another distribution. The KL divergence is one of the statistical distance measures, which are mostly not metrics and not have to be symmetric. The KL divergence for two discrete distributions, Q_1 and Q_2, is defined as follows:

$$D_{KL}(Q_1 \mid\mid Q_2) = \sum_i Q_1(i) \log \left(\frac{Q_1(i)}{Q_2(i)} \right). \tag{C.2}$$

Similarly, the KL divergence for two continuous distributions, $q_1(x)$ and $q_2(x)$, is defined as

$$D_{KL}(q_1(x) \mid\mid q_2(x)) = \int_{-\infty}^{\infty} q_1(x) \log \left(\frac{q_1(x)}{q_2(x)} \right) dx. \tag{C.3}$$

For more detailed properties of the KL divergence, see Kullback (1968).